

Simulation of high-latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e) 1: Experiment description and summary intercomparisons

Laura C. Bowling^{a,*}, Dennis P. Lettenmaier^a, Bart Nijssen^{a,1}, L. Phil Graham^b,
Douglas B. Clark^c, Mustapha El Maayar^d, Richard Essery^e, Sven Goers^f,
Yeugeniy M. Gusev^g, Florence Habets^h, Bart van den Hurkⁱ, Jiming Jin^j,
Daniel Kahan^k, Dag Lohmann^l, Xieyao Ma^m, Sarith Mahanamaⁿ, David Mocko^o,
Olga Nasonova^g, Guo-Yue Niu^j, Patrick Samuelsson^{b,p}, Andrey B. Shmakin^q,
Kumiko Takata^m, Diana Verseghy^r, Pedro Viterbo^s, Youlong Xia^{t,2},
Yongkang Xue^k, Zong-Liang Yang^u

^aDepartment of Civil and Environmental Engineering, University of Washington Box 352700, Seattle, WA 98195-2700, USA

^bSwedish Meteorological and Hydrological Institute, Sweden

^cCentre for Ecology and Hydrology, Wallingford, UK

^dCanadian Forest Service, Northern Forestry Centre, Edmonton AB, Canada

^eMet Office, UK

^fGKSS Research Center, Geesthacht, Germany

^gInstitute of Water Problems, Russian Academy of Sciences, Russian Federation

^hNational Center for Atmospheric Research/CNRM, Toulouse, France

ⁱRoyal Netherlands Meteorological Institute, De Bilt, The Netherlands

^jDepartment of Hydrology and Water Resources, University of Arizona, Tucson, AZ, USA

^kUniversity of California Los Angeles, Torrance, CA, USA

^lNCEP Environmental Modeling Center (NOAA/NWS), USA

^mFrontier Research System for Global Change, Japan

ⁿHydrological Sciences Branch, NASA Goddard Space Flight Center and GEST/University of Maryland Baltimore County, USA

^oNASA Goddard Space Flight Center and SAIC/General Sciences Operation, USA

^pDepartment of Meteorology, Stockholm University, Stockholm, Sweden

^qInstitute of Geography, Russian Academy of Sciences, Moscow, Russian Federation

^rClimate Research Branch, Meteorological Service of Canada, Canada

^sEuropean Centre for Medium-Range Weather Forecasts, Reading, UK

^tMacquarie University, North Ryde, NSW, Australia

^uDepartment of Geological Sciences, University of Texas at Austin, Austin, TX, USA

Received 17 November 2001; received in revised form 28 January 2002; accepted 3 June 2002

* Corresponding author. Fax: +1-206-616-6274.

E-mail address: lxh@hydro.washington.edu (L.C. Bowling).

¹ Present address: Department of Civil Engineering and Hydrology and Water Resources, University of Arizona, Tucson, AZ, USA.

² Present address: Institute for Geophysics, University of Texas at Austin, USA.

Abstract

Twenty-one land-surface schemes (LSSs) participated in the Project for Intercomparison of Land-surface Parameterizations (PILPS) Phase 2(e) experiment, which used data from the Torne–Kalix Rivers in northern Scandinavia. Atmospheric forcing data (precipitation, air temperature, specific humidity, wind speed, downward shortwave and longwave radiation) for a 20-year period (1979–1998) were provided to the 21 participating modeling groups for 218 $1/4^\circ$ grid cells that represented the study domain. The first decade (1979–1988) of the period was used for model spin-up. The quality of meteorologic forcing variables is of particular concern in high-latitude experiments and the quality of the gridded dataset was assessed to the extent possible. The lack of sub-daily precipitation, underestimation of true precipitation and the necessity to estimate incoming solar radiation were the primary data concerns for this study. The results from two of the three types of runs are analyzed in this, the first of a three-part paper: (1) calibration–validation runs—calibration of model parameters using observed streamflow was allowed for two small catchments (570 and 1300 km²), and parameters were then transferred to two other catchments of roughly similar size (2600 and 1500 km²) to assess the ability of models to represent ungauged areas elsewhere; and 2) reruns—using revised forcing data (to resolve problems with apparent underestimation of solar radiation of approximately 36%, and certain other problems with surface wind in the original forcing data). Model results for the period 1989–1998 are used to evaluate the performance of the participating land-surface schemes in a context that allows exploration of their ability to capture key processes spatially. In general, the experiment demonstrated that many of the LSSs are able to capture the limitations imposed on annual latent heat by the small net radiation available in this high-latitude environment. Simulated annual average net radiation varied between 16 and 40 W/m² for the 21 models, and latent heat varied between 18 and 36 W/m². Among-model differences in winter latent heat due to the treatment of aerodynamic resistance appear to be at least as important as those attributable to the treatment of canopy interception. In many models, the small annual net radiation forced negative sensible heat on average, which varied among the models between -11 and 9 W/m². Even though the largest evaporation rates occur in the summer (June, July and August), model-predicted snow sublimation in winter has proportionately more influence on differences in annual runoff volume among the models. A calibration experiment for four small sub-catchments of the Torne–Kalix basin showed that model parameters that are typically adjusted during calibration, those that control storage of moisture in the soil column or on the land surface via ponding, influence the seasonal distribution of runoff, but have relatively little impact on annual runoff ratios. Similarly, there was no relationship between annual runoff ratios and the proportion of surface and subsurface discharge for the basin as a whole.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Arctic regions; modelling; hydrology; atmosphere

1. Introduction

Climate models predict that temperature changes associated with global warming may be greatest at high latitudes (Giorgi et al., 2001). The temperature sensitivity of high-latitude land areas is due in part to climate feedbacks, such as the reduction of snow extent and the resulting decrease in wintertime land-surface albedo that would likely accompany a warmer climate, leading to a subsequent increase in absorbed radiation (Giorgi et al., 2001). In addition, Arctic freshwater flux, along with sea-ice formation, controls the salinity of the polar ocean and thus the thermohaline circulation of the world ocean (Broecker, 1997). Coupled ice-ocean models indicate that decadal trends in both arctic precipitation and river runoff have the

potential to exert broad-scale impacts on the arctic sea ice regime (Weatherly and Walsh, 1996) and to excite decadal and interdecadal oscillations of ocean circulation (Weaver et al., 1991). Steele et al. (1996) found direct linkages between the variability of freshwater in the Arctic Ocean and transport rates out of Fram Strait, located between Greenland and the islands of Svalbard, relative to fluxes through the Canadian Archipelago. A large increase in freshwater flux to the convective gyres in the Greenland and Iceland seas through Fram Strait could have major consequences for convection and global thermohaline ocean circulation (Aagaard and Carmack, 1989).

Runoff from the land surface represents the single largest input of freshwater to the Arctic Ocean. By one estimate, up to 41% of the discharge from land areas

into the Arctic Ocean is derived from ungauged basins and islands (WCRP, 1996). Although extrapolation of observed streamflow to ungauged areas allows bulk estimates of land-surface freshwater fluxes, the temporal and spatial variability of runoff response from these areas is unknown. For this reason, a 1998 workshop on the status and directions of the Arctic runoff database hosted by the World Climate Research Programme's (WCRP) Arctic Climate System Study (ACSYS) (IAPO, 1998) concluded that hydrological modeling offered the most plausible method of estimating runoff from the ungauged areas. Similarly, the WCRP's Global Energy and Water cycle EXperiment (GEWEX) is concerned with the water and energy balance of two high-latitude continental study areas—the Mackenzie River basin, which constitutes the Mackenzie GEWEX Study (MAGS) and the Lena River basin, which is one of the GEWEX Asian Monsoon Experiment (GAME) study areas.

Coupled land–atmosphere–ocean models are important tools for studies of the interaction of land, atmosphere and ocean processes as they affect climate. Traditionally, however, land-surface schemes (LSSs) used in numerical weather prediction and climate models have treated high-latitude processes quite crudely. More recently, recognition of the sensitivity of high-latitude land areas to climate change, and the resultant feedbacks on climate such as those outlined above, has motivated research focused on improving the representation of cold season processes in LSSs. Many modeling groups have changed their representation of snow (e.g., Shmakin, 1998; Jin et al., 1999a; Sud and Mocko, 1999; Viterbo and Betts, 1999; van den Hurk et al., 2000; Boone and Etchevers, 2001; Sun and Xue, 2001; Niu and Yang, in preparation) and frozen soil processes (e.g., Cherkauer and Lettenmaier, 1999; Viterbo et al., 1999; Boone et al., 2000; van den Hurk et al., 2000). Lynch et al. (1998) found that the largest differences between the BATS and LSM LSSs were due to snow albedo and the treatment of partial snow coverage. Mocko and Sud (2001) found that improvements to the SSiB snow model and infiltration scheme improved the simulated snowmelt timing and runoff. In a comparison of three snow models (SNTHERM, SAST and BATS), Jin et al. (1999b) found biases in the representation of surface temperature and energy fluxes in the BATS model, which they attributed to neglecting the effects of liquid water

in the snow pack. They obtained results comparable to the more complex SNTHERM model by adding two thin upper layers to the SAST snow model, to better represent near-surface variations in heat conduction and melt water heat transport (Jin et al., 1999b). In a comparison of three LSSs (BATS, LSM and CLASS), Tilley and Lynch (1998) found a wide range in the ability of the models to capture the seasonal cycle of ground temperatures, which they attributed to the number of soil layers represented. Takata and Kimoto (2000) found that including the effects of soil freezing (latent heat of fusion and impermeability of frozen ground) in an atmospheric general circulation model (CCSR/NIES) leads to higher simulated summer surface temperatures in the middle and high latitudes. This results in increased simulated precipitation in Southeast Asia. Precipitation is lower over Eurasia and North America due to the limitation on evaporation of decreased soil water (Takata and Kimoto, 2000). The effects of lakes and wetlands on surface storage and the timing of arctic runoff, and the effects of sublimation from blowing snow on the surface water and energy balances, are largely neglected in most current generation LSSs.

The importance of Arctic land-surface dynamics to global climate prompted a closer evaluation of the performance of high-latitude land-surface representations within land–atmosphere models used for numerical weather and climate prediction. A small working group was established by ACSYS in August 1998 for the purposes of planning an arctic hydrology model intercomparison project. This working group recognized that an intercomparison of hydrological models in northern environments could benefit the goals of both GEWEX and ACSYS by quantifying the accuracy with which current land schemes represent high-latitude land processes, by providing information about pathways for model improvements, and by providing information about the accuracy with which land schemes can be used to estimate runoff from ungauged areas draining to the Arctic Ocean (WCRP, 1999a). The objective of the intercomparison project described in this, and companion papers (Nijssen et al., 2003-this issue; Bowling et al., 2003-this issue) is to evaluate the performance of uncoupled land-surface parameterizations in high latitudes, in a context that allows evaluation of their ability to capture key processes spatially. These include snow accumulation and

ablation, soil freeze/thaw and permafrost, and the existence of large seasonally frozen lakes and wetlands.

The intercomparison study outlined in [WCRP \(1999b\)](#) was undertaken as a joint experiment of ACSYS and GEWEX under the auspices of the GEWEX Project for Intercomparison of Land-surface Parameterization Schemes (PILPS). The family of PILPS Phase 2 experiments involves off-line testing of LSSs using observed forcing data ([Henderson-Sellers et al., 1995](#)). The organization of the experiment reported in this three-part series, and designated PILPS Phase 2(e), follows the classic PILPS Local Off-Line framework outlined by [Henderson-Sellers et al. \(1993, 1995\)](#). An important aspect of Phase 2(e) is the use of a consistent data exchange format scheme developed in cooperation with the Assistance for Land-surface Modeling Activities (ALMA) action group of GLASS ([Polcher et al., 2000](#)). Phase 2(e) follows in the tradition of the first spatial experiment, Phase 2(c), over the Arkansas–Red River basin in the United States ([Wood et al., 1998](#)) and Phase 2(d), a previous cold regions experiment conducted at a grassland site in Valdai, Russia ([Schlosser et al., 1997](#)).

The Phase 2(e) intercomparison described here utilizes data from the Torne and Kalix River system in northern Scandinavia. Although these rivers do not drain to the Arctic Ocean, they do lie within an arctic environment (characterized by annual average net radiation near zero), and were chosen as the initial test location for Phase 2(e) to take advantage of the relatively dense and lengthy observation records obtained from networks maintained by the Swedish and Finnish Meteorological and Hydrological Institutes. Even so, the data available for this remote and extreme environment is significantly less than in previous PILPS intercomparison experiments carried out at lower latitudes. The strategy for Phase 2(e) was to test the performance of the LSSs and hydrology models as they are usually implemented. Therefore, rather than providing the participants with detailed vegetation and soil data which in any event were not available, participants were encouraged to use in-house methods to derive the required model parameters.

Twenty-one LSSs participated in the Torne–Kalix experiment (see [Table 1](#)). This paper describes the experiment design, and summarizes and evaluates the model forcing data sets that were provided to the

Table 1
List of participating models of PILPS Phase 2(e)

Model	Contact	Calibrate sub-basins?	Apply to basin? ^a	Input time step (min)	Output time step (min)
CHASM	Y. Xia, A. Pitman, A. Henderson-Sellers	No	No	60	60
CLASS	D. Verseghy	Yes	No	30	60
ECMWF	P. Viterbo, B. van den Hurk	No	No	30	60
HY-SSiB	D. Mocko, Y. Sud	Yes	Yes	60	60
IBIS	M. El Maayar, J. Foley	No	No	60	60
IHAS	X. Ma	Yes	No	60	60
ISBA	F. Habets, A. Boone, J. Noilhan	Yes	Yes	5	5
MATSIRO	K. Takata	Yes	Yes	60	60
MECMWF	B. van den Hurk, P. Viterbo	Yes	Yes	30	60
MOSES	R. Essery	No	No	60	60
MOSES-CEH	D. Clark, E. Blyth	No	No	60	60
NOAH	D. Lohmann, K. Mitchell	No	No	15	60
NSIPP	R. Koster, S. Mahanama	Yes	Yes	20	60
RCA	P. Samuelsson	No	No	15	15
SAST	J. Jin, X. Gao, S. Sorooshian	Yes	Yes	60	60
SEWAB	S. Goers, H. Mengelkamp, M. Klein	Yes	Yes	30	60
SPONSOR	A. Shmakin	No	No	60	60
SSiB	Y. Xue, D. Kahan	Yes	Yes	60	60
SWAP	Y. Gusev, O. Nasonova	No	No	60	60
VIC	L. Bowling, D. Lettenmaier	Yes	Yes	60	60
VISA	Z. Yang, G. Niu	Yes	Yes	60	60

^a Was calibration knowledge from the Övre Abiskojokk and the Övre Lansjärv sub-basins used to adjust parameters for the entire basin?

participants. [Nijssen et al. \(2003-this issue\)](#) assess the ability of the schemes to capture the processes controlling surface energy and moisture fluxes in this high-latitude environment through intercomparison of simulated results and observations. [Bowling et al. \(2003-this issue\)](#) report on sensitivity analyses conducted through comparison of the various models with output from simplified equivalent bucket models.

2. Background

The combined size of the Torne and Kalix River Basins is approximately 58,000 km². Approximately 75% of the area lies in Sweden, 25% in Finland, and less than 1% in Norway (Fig. 1). The main stem of the Torne River forms the border between Sweden and Finland and flows into the Gulf of Bothnia, an arm of the Baltic Sea. The Torne–Kalix basin is therefore part of the BALTEX (Baltic Sea Experiment) domain, a GEWEX project to determine the energy and water

budgets of the Baltic Sea and related river basins. A natural bifurcation in the drainage network diverts on average 22% of the Torne River runoff into the adjacent Kalix River basin (Fig. 2) above the Junosuando/Tarendo station (Carlsson, 1999). The two rivers are therefore treated as one hydrologic unit for the purposes of this experiment. The bifurcation is gauged, so the effect of this diversion can be accounted for in the routing of simulated runoff. Near the mouth of the Torne at Karunki, the tributary Liankanjoki draws off another 6% of the discharge

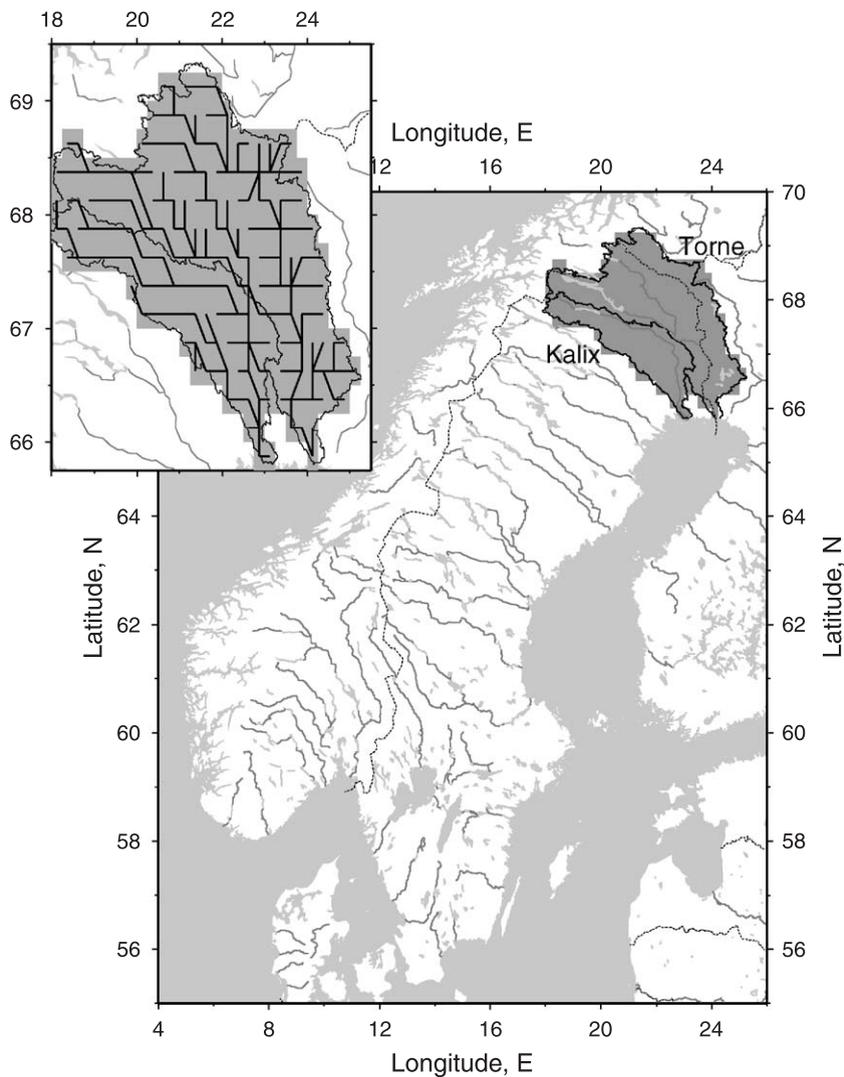


Fig. 1. Location within Scandinavia of the Torne–Kalix River basin and the 218 computational 1/4° latitude/longitude grid cells.

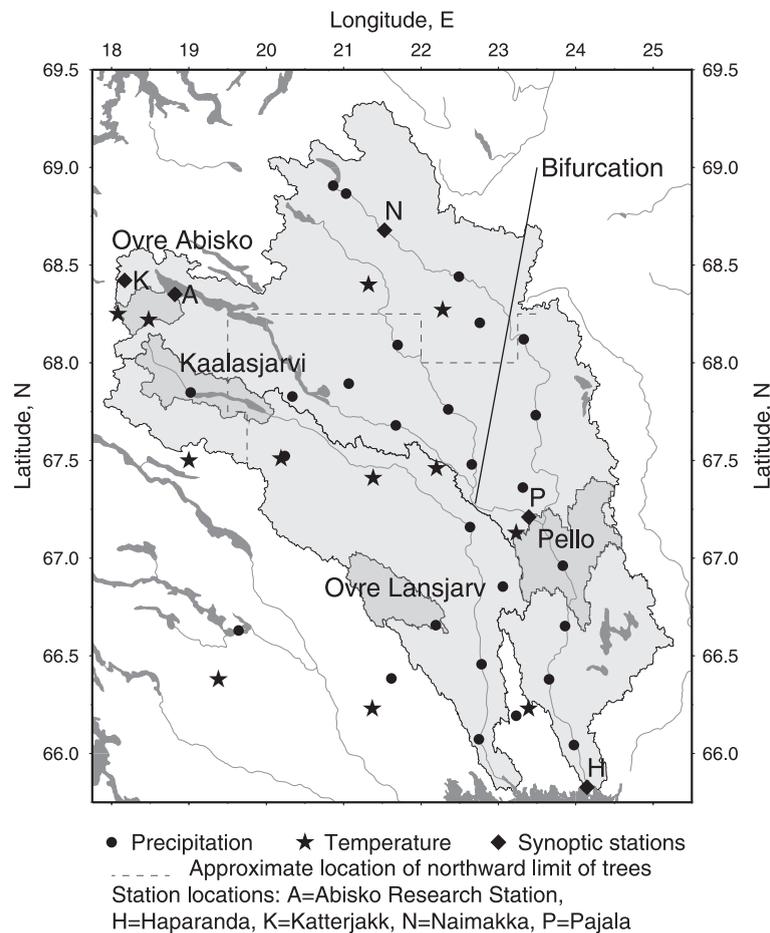


Fig. 2. The Torne–Kalix River system and observation station locations.

on average (Carlsson, 1999). Discharge at Karunki, above the diversion was used for model comparison purposes. Both the Torne and the Kalix Rivers are essentially unregulated.

Precipitation within the study area is characterized by an increasing gradient with elevation to the north and west that reaches a maximum of approximately 1600 mm/year. The mountains in the northwest make up 7–8% of the total drainage area and include Mount Kebnekaise, the highest point in Sweden (maximum elevation 2117 m). There is also a regional maximum in precipitation on the order of 700 mm within approximately 75 km of the Baltic coast. The central third of the basin lies between 200 and 500 m above sea level. Precipitation in this area averages between 500 and 600 mm annually.

Annual average temperature varies between $+1\text{ }^{\circ}\text{C}$ in the south and $-3.5\text{ }^{\circ}\text{C}$ in the north, and the average number of days with snow cover ranges from 175 to 225 days (Carlsson, 1999). Snow processes therefore affect the entire basin and the runoff regime is characterized by a seasonal maximum following snowmelt. Much of the runoff is produced in the high-latitude areas above the tree line where precipitation is the highest.

Vegetation consists primarily of needle leaf boreal forests interspersed with large open areas of low relief mire and bog throughout the central portion of the basin (Carlsson, 1999). Birch trees also comprise 10–20% of the vegetation in central and coastal areas. With increasing elevation and latitude, land cover transitions from forest to bare rock, moss and lichens (Carlsson, 1999).

Surficial soil deposits within the Torne and Kalix basins consist predominantly of glaciofluvial sediments, till and peat (Freden, 1994). Soils are thin in the mountains and consist of exposed rock or areas with a very thin overburden that has been affected by frost processes to different degrees. The Veiki moraine, characterized by plateau-like moraine hills with numerous circular-formed lakes, is found in a wide belt across the central portion of both basins. Peat is prevalent in the soils of the moraine and these deposits are often quite thick (Freden, 1994). The coastal area (below approximately 100 m) is dominated by sandy till, with glaciofluvial and fine-grained sediment deposits in the valleys (Freden, 1994).

The Swedish Meteorological and Hydrological Institute (SMHI) has an extensive observation network for hydrometeorological data within the Torne–Kalix Rivers that dates back several decades (see, e.g., Fig. 2). The SMHI data constitute the primary source of forcing data for the models, as described below.

3. Experimental design

3.1. Data sources

3.1.1. Land-surface characteristics

The land-surface characteristics provided to the participants included both spatially distributed (cell-specific) data and look-up table parameters. No attempt was made to specify all parameters required by each model. A basic set of characteristics and descriptive information was provided from which each modeling group could derive the necessary parameters. Free calibration of model parameters was only permitted (and only possible) for the two calibration basins, Övre Lansjärv and Övre Abisko-jokk, for which observed discharge was provided to the participants. Spatial parameters were derived from the following databases.

- *Basin masks*—Basin boundaries for the Torne and Kalix River Basins and the four calibration/validation sub-basins were provided by SMHI. The fractional area draining to the basin mouth was computed for each 1/4° grid cell from these boundaries using Arc/Info GIS software.

- *Topography*—The mean elevation of each 1/4° grid cell was calculated based on the Global 30 Arc

Second Elevation Data Set (GTOPO30) distributed by the EROS Data Center Distributed Active Archive Center (EDC DAAC), located at the U.S. Geological Survey.

- *Soils*—Soil textural information and soil bulk densities for the top meter of soil were obtained for each 1/4° grid cell using the SoilProgram (Carter and Scholes, 1999), which combines the 5-min FAO-UNESCO digital soil map of the world (FAO, 1995) with the WISE pedon database (Batjes, 1995). Soil hydraulic properties were specified using relationships based on USDA soil texture (Clapp and Hornberger, 1978; Rawls et al., 1982; Cosby et al., 1984). All participants were requested to use one of these three parameter sets for porosity, saturated soil matric potential, saturated soil hydraulic conductivity and the Clapp-Hornberger ‘B’ parameter used to calculate the unsaturated hydraulic conductivity. The moisture contents at field capacity and wilting point were determined by calculating the moisture retention at a matric pressure of –33 and –1500 kPa, respectively. The list of provided parameters was selected based on the protocol established by the Land Data Assimilation System (LDAS) project (Mitchell et al., 1999), which utilizes procedures similar to those used in the Phase 2(c) project (Wood et al., 1998).

- *Vegetation*—Vegetation types were determined using the 1-km AVHRR-based global land classification of Hansen et al. (2000). This classification uses 14 land-surface classes, including water and bare soil, of which 10 are found in the Torne–Kalix basin. The urban land class was combined with the bare ground class, resulting in a total of nine classes. Vegetation height and albedo were taken from a variety of sources.

- *LAI*—Leaf area index (LAI) values were based on the 15-min global data set of Myneni et al. (1997), as processed by Nijssen et al. (2001). Myneni et al. (1997) calculated monthly LAI values by relating AVHRR-based Normalized Difference Vegetation Indices (NDVI) to LAI using a three-dimensional radiative model. Nijssen et al. (2001) mapped the 15-min LAI values to individual vegetation types by identifying 15-min grid cells of homogeneous vegetation from the 1-km dataset of Hansen et al. (2000). For this experiment, the monthly LAI values for all

homogeneous grid cells within the Torne–Kalix basin for a given vegetation type were averaged.

3.1.2. Meteorological variables

The following hourly surface meteorological variables were provided to participants for the period 1979–1998: rain and snow precipitation, air temperature, surface pressure, specific humidity, downward longwave and shortwave radiation, wind speed and cloud cover fraction. Station data were provided by SMHI. A 1° dataset of several key variables optimally interpolated from station data (Mueller, 1998) was also provided by SMHI for the BALTEX domain.

The quality of meteorologic forcing variables is of particular concern in high-latitude experiments due to the sparseness of observation stations and the difficulty of measurement and maintenance in remote locations subject to harsh winter conditions. Low station density and a lack of sub-daily precipitation data necessitated adjusting the daily precipitation gridded from station data for elevation, and statistically disaggregating daily precipitation to hourly. The predominance of frozen precipitation in this basin results in a high-probability of under-observation of precipitation, so the station data were corrected for gauge undercatch, as discussed in Section 4.4. Estimation of shortwave radiation was also problematic, as described in Section 4.3, and resulted in the distribution of a second forcing dataset mid-way through the experiment.

Gridded temperature data at $1/4^\circ$ resolution were generated by interpolation of the station data and lapsed with elevation, as described in Section 4.2. Cloud cover, relative humidity and surface pressure were observed at a smaller number (five) of synoptic stations in and around the Torne–Kalix basin for the entire period (1979–1998) and at six automatic stations for the period 1990–1996. For these variables, a $1/4^\circ$ dataset was therefore generated by linear interpolation of the data of Mueller (1998), see Section 4.1. Longwave radiation was calculated using gridded air temperature, humidity and cloud cover, as described in Section 4.3. Wind speeds were taken from the daily surface wind fields (10 m) of the NCEP/NCAR reanalysis (Kalnay et al., 1996), linearly interpolated in space. Following interpolation and gridding of all data, the data sets were screened for

pathological values in accordance with the ALMA protocols.

3.2. Experiment set-up

For the purposes of the experiment, the 58,000- km^2 Torne and Kalix River Basins were represented together by 218 $1/4^\circ$ computational grid boxes (Fig. 1). Three types of model runs were requested from the participants:

- (1) Calibration and validation runs. Model forcing data and streamflow observations were provided to each modeling group for two sub-catchments of the Torne–Kalix system, the Övre Abiskojojk and the Övre Lansjärv (see Fig. 2). These catchments were selected to represent a range in elevation and in characteristics of vegetation cover. For two similarly selected validation catchments (Pello and Kaalasjärvi), only model forcing data were provided to the participants. Parameters were to be transferred to the validation catchments, and to the basin as a whole, using methods of the participants' choice. This design mimics a similar experiment conducted as part of PILPS Phase 2(c) (Wood et al., 1998).
- (2) Base-runs. For these runs, the modelers simulated the surface energy and moisture fluxes for each of the 218 $1/4^\circ$ grid cells. Core diagnostics performed by the organizers on the submitted model output were intended to evaluate the representation of cold region processes, such as snow accumulation and ablation and soil freezing and thawing, at the scale of the Torne–Kalix River basin.
- (3) Reruns. Due to problems with the original shortwave radiation and wind speed fields, updated forcing data were provided to the participants for a rerun of the base-run experiment. The basinwide mean annual incoming solar radiation increased 36% from 66.2 to 90.2 W/m^2 and the basinwide mean annual wind speed increased from 1.2 to 3.1 m/s for the model reruns.

All results for experiments 2 and 3 were screened to assure that water and energy were balanced in their annual means. These consistency checks were set at 3 mm for the water balance and 3 W/m^2 for the energy

balance, as adopted by earlier PILPS experiments. Only the results of experiments 1 and 3 are addressed here. A workshop was held at the University of Washington following experiment 2, where participants saw their results compared to observations and the other models for the first time. Approximately one half of the models submitted their results for experiment 3 after the workshop and these simulations may have been influenced in part by knowledge of relative model performance. In particular, the RCA model had unrealistically large estimates of wintertime latent heat in the results presented at the workshop. A reduction of the snow surface roughness resulted in a 40% reduction in predicted annual latent heat for the final submission (Samuelsson et al., 2003-this issue).

The revised forcing data created for experiment 3 was also used for the experiment 1 calibration and validation runs discussed here. As noted in Section 3.1.2, for all three runs hourly forcing data were supplied for the 20-year time period, 1979–1998. The period 1989–1998 was used for the intercomparison. Some participants disaggregated the forcing data to conduct the simulations at a finer temporal scale.

4. Data preparation

4.1. Humidity and pressure

Specific humidity and air pressure were derived from the 1° dataset of Mueller (1998). Air pressure at sea level was linearly interpolated to the center of each $1/4^\circ$ grid cell and converted to surface pressure using the mean grid cell elevation. Relative humidity was converted to specific humidity using the gridded air pressure and temperature, and was linearly interpolated to the grid cell centers. To verify that no bias was introduced in the conversion and gridding of pressure and humidity, relative humidity was back calculated from the final $1/4^\circ$ fields and compared with the three-hourly relative humidity observed at four synoptic stations located within the Torne–Kalix domain, for the entire 1979–1998 time period. As summarized in Table 2 and illustrated in Fig. 3, the total bias does not exceed $\pm 2\%$ for any one station on an annual or seasonal basis, and takes on both positive and negative values, depending on the station and/or time period.

Table 2

Bias with respect to observed relative humidity

Station	Winter		Summer		Total	
	MSE	Bias (%)	MSE	Bias (%)	MSE	Bias (%)
Haparanda	34.2	−0.4	72.8	−1.4	53.5	0.9
Katterjåkk	59.4	2.3	72.4	1.6	65.9	2.0
Naimakka	37.5	0.5	55.5	0.4	46.5	0.5
Pajala	38.9	−1.8	77.8	−2.1	58.1	−1.9

4.2. Air temperature

Daily 2-m air temperature for the study period (1979–1998) was available from 19 SMHI stations within the Torne–Kalix basin. Station data were interpolated to the mid-point of each $1/4^\circ$ grid cell using the SYMAP interpolation scheme (Shepard, 1984) and lapsed to the elevation of the cell mid-point using a constant lapse rate of $-0.0065^\circ\text{C}/\text{m}$. The daily gridded data were disaggregated to a three-hour time step using anomalies from the Mueller (1998) dataset to create a three-hourly time series while preserving the daily mean of the $1/4^\circ$ cell. Air temperatures were linearly interpolated from the three-hourly instantaneous values to produce a mean hourly time series. The constructed hourly time series were compared to three-hourly observations at six stations for 1996–1998, as illustrated in Fig. 4. The comparison shows that there is no clear bias in the estimations over the entire range of observed temperatures, however, there is a tendency for smoothing of daily extreme temperatures, and in particular a tendency to predict temperatures that are too warm when observed temperatures are extremely low. This should not have a large effect on the energy state of the cold mid-winter snow pack relative to the between-model differences.

4.3. Shortwave and longwave radiation

4.3.1. Shortwave radiation

Incoming solar radiation was originally calculated using the method of TVA (1972) with standard parameters derived from look-up tables. Cloud cover was taken from the Mueller (1998) dataset, linearly interpolated to $1/4^\circ$. After the experiment was underway, it became clear that this method produced a negative bias in the incoming shortwave radiation.

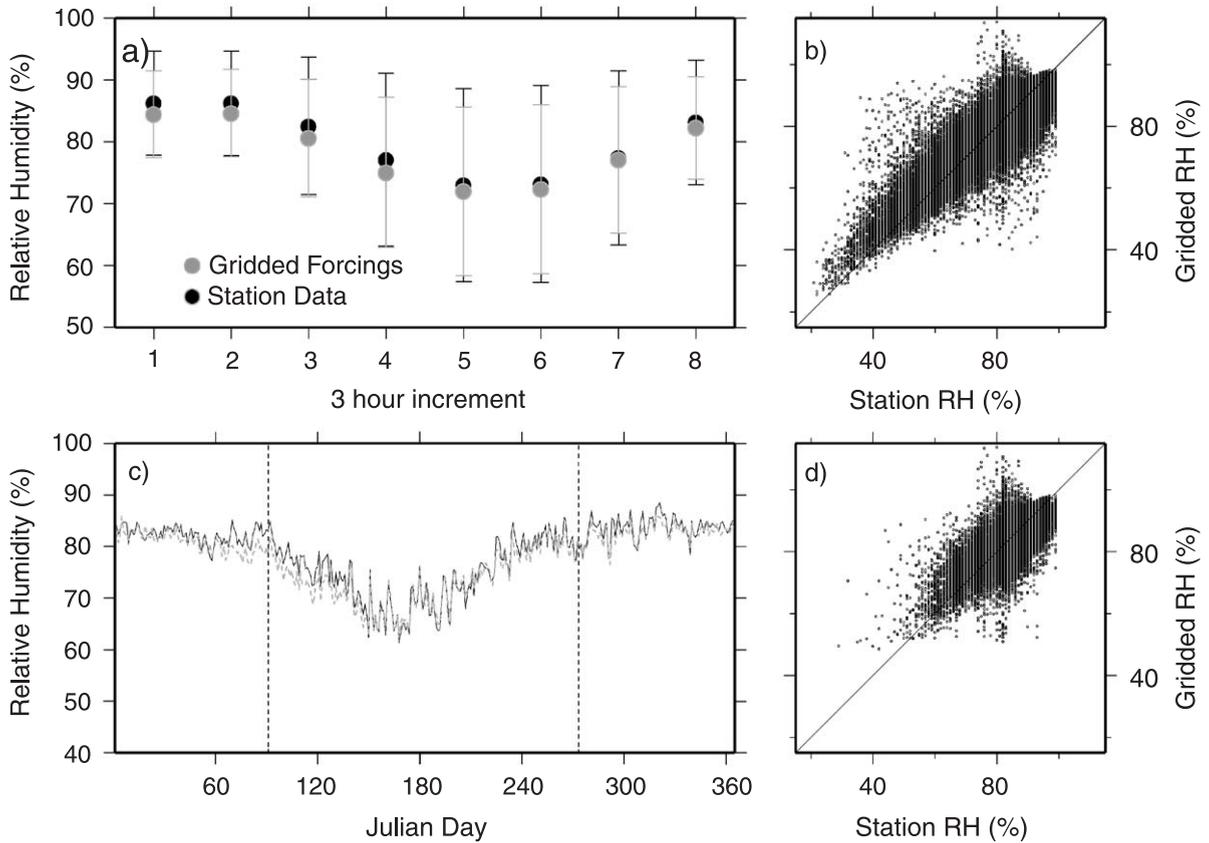


Fig. 3. Observed relative humidity at the Naimakka station vs. gridded estimates: (a) mean diurnal cycle, (b) summer observations, (c) mean annual cycle and (d) winter observations.

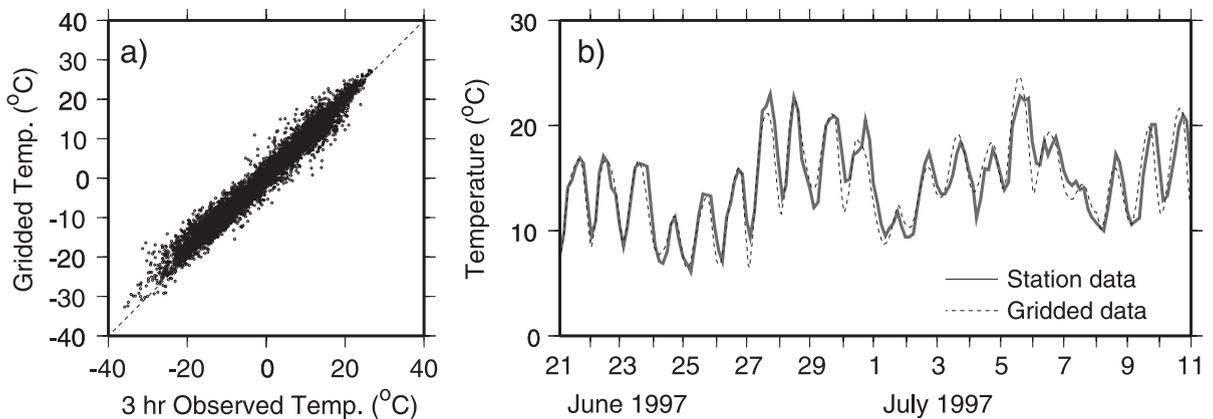


Fig. 4. Gridded temperature versus observations at Paharova: (a) three-hourly observations and (b) time series of three-hourly observations.

Daily average solar radiation for the period January 1969 through December 1993 at Kiruna in the Torne–Kalix basin was obtained from the World Radiation Data Centre (WRDC) located at the Main Geophysical Observatory in St. Petersburg, Russia. As indicated in Fig. 5a, the PILPS incoming shortwave radiation computed using the TVA (1972) method has a low bias in comparison with the Kiruna data. Fig. 5b and c shows the results of the Eagleson (1970) method, combined with the same cloud cover, to estimate clear-sky radiation, as follows:

$$I_c = I_o e^{-na_1 m} \quad (1)$$

where I_c is the clear-sky radiation and I_o is the solar insolation incident at the top of the atmosphere. The molecular scattering coefficient, a_1 , and the optical air

mass, m , are both calculated as a function of solar altitude. The turbidity factor for air, n , was calibrated based on the Kiruna station data to a value of 1.25. Solar radiation was also calculated at Luleå, a station on the coast just south of the Kalix basin, using the same turbidity factor of 1.25. Bias and mean squared error (MSE) for the revised PILPS dataset was calculated with respect to the daily time series at Kiruna and Luleå for the period (1979–1993). As indicated in Table 3, calculations of solar radiation using the turbidity factor estimated for Kiruna appear to be reasonable at Luleå, although the results for the summer months are not as good as at Kiruna.

The change in shortwave radiation between the base runs and reruns is substantial. It may be possible to quantify the effect of this change if the effect of the

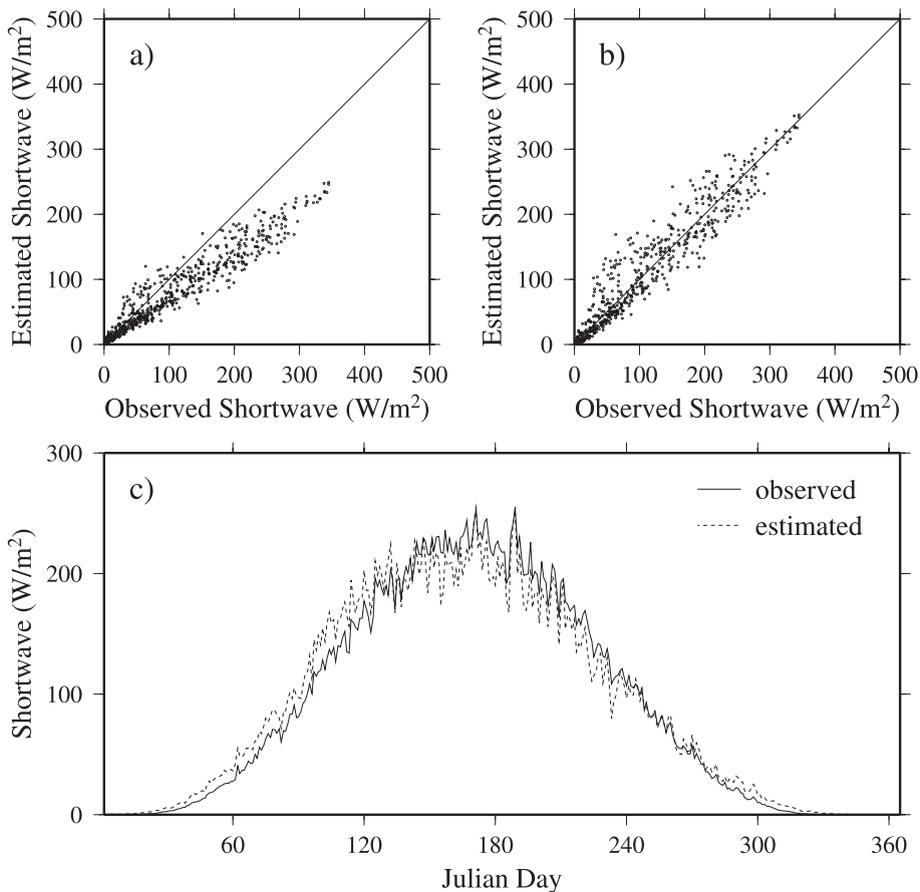


Fig. 5. Observed daily average solar insolation at Kiruna: (a) versus the Phase 2(e) base-run forcing data, (b) versus the Phase 2(e) rerun forcing data and (c) mean annual cycle of the rerun forcing data.

Table 3
Bias with respect to observed solar radiation

Station	Winter		Summer		Total	
	MSE	Bias	MSE	Bias	MSE	Bias
	(W/m ²)		(W/m ²)		(W/m ²)	
Kiruna—rerun	96.1	5.4	1180.8	−4.0	678.0	0.34
Luleå—rerun	126.4	0.9	3863.8	−21.2	2136.8	−10.9
Kiruna—base run	147.1	6.4	3381.1	42.3	1878.1	25.6
Luleå—base run	166.4	3.1	4004.4	32.7	2233.1	19.1

increase in shortwave radiation can be segregated from that of the increase in wind speed and the precipitation disaggregation. In most cases, the precipitation disaggregation is not expected to have much influence on the surface energy partitioning, with the exception of a few models that reported instabilities in latent heat under the constant precipitation apportionment, as discussed in Section 4.4.

4.3.2. Longwave radiation

Downward longwave radiation for each grid cell was calculated as a function of the gridded two-m air temperature, using an equation for emissivity (E_a) from TVA (1972):

$$E_a = 0.74 + 0.0049e_o \quad (2)$$

where e_o is the vapor pressure in millibars. The calculated radiation was corrected for cloud cover by the method of TVA (1972) using the observed

cloud cover taken from the Mueller (1998) dataset, linearly interpolated to 1/4°.

The International Satellite Land Surface Climatology Project (ISLSCP) Initiative I data set included global data for the years 1987–1988 at a resolution of 1° latitude by 1° longitude (Sellers et al., 1995). Longwave radiation data were estimated from thermal remote sensing. A comparison of the PILPS longwave data, averaged over the 208 grid cells between 18 to 25°E and 66 to 69°N with the 21 corresponding ISLSCP grid cells indicates that for the 2 years, 1987–1988, the PILPS data are on average higher by about 9 W/m² than the ISLSCP data, as illustrated in Fig. 6.

Because there are also errors inherent in the method of estimation used by ISLSCP, this difference is not necessarily an indication of bias in the PILPS estimation technique. Observations of incoming longwave radiation were available for one station, the Abisko Scientific Research Station (68°21' N, 18°49' E), in the headwaters of the Torne River (see Fig. 2). Downward longwave radiation was calculated for this site in Eq. (2) using daily air temperature, humidity and pressure for the period 1989–1998 and compared to observed (shown graphically in Fig. 7). On average, the calculated longwave radiation was 9 W/m² (3%) lower than observed. Longwave radiation from the gridded PILPS dataset was 6 W/m² lower than observed on average. These calculations tend to con-

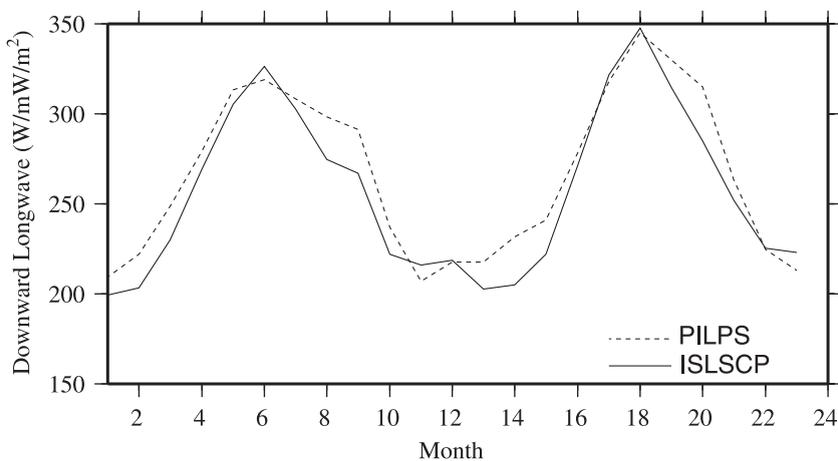


Fig. 6. Average monthly downward longwave radiation (1987–1988) from the Phase 2(e) dataset, averaged over 208 grid cells, versus the International Satellite Land Surface Climatology Project (ISLSCP) longwave radiation between 18 to 25°E and 66 to 69°N.

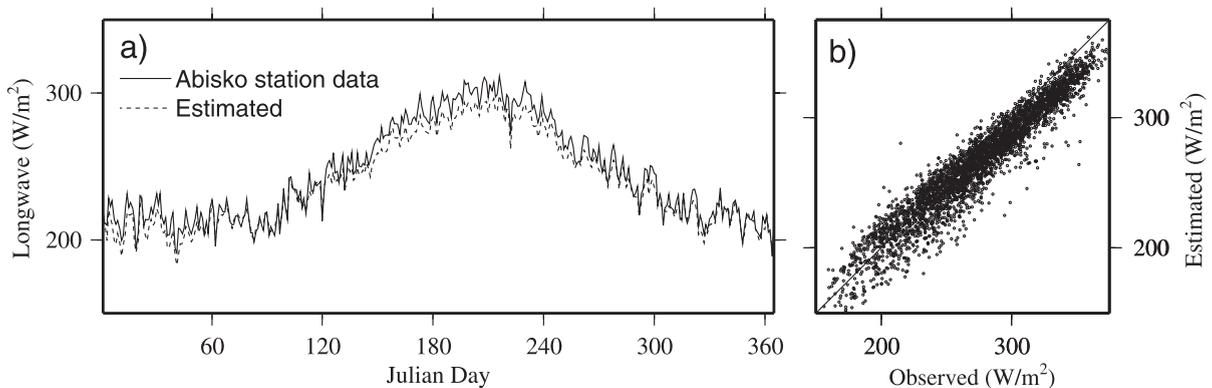


Fig. 7. Daily downward longwave radiation estimated using the method of the Tennessee Valley Authority (TVA 1972) versus observations at Abisko, in northern Sweden for 1989–1998. (a) Mean annual cycle and (b) mean daily versus daily observed.

tradict the magnitude and sign of bias compared to the ISLSCP data and we conclude that there is no clear tendency toward over or underprediction in the longwave radiation forcing data.

Schlosser et al. (2000) conducted a sensitivity experiment with 21 LSSs using the Brutsaert (1975) and Idso (1981) methods of estimation of longwave radiation as part of the Phase 2(d) experiment for a grassland site in Valdai, Russia. An imposed decrease in winter longwave radiation of 30–40 W/m^2 resulted in an increase in average snow water equivalent (SWE) on the order of 40 to 65 mm (approximately 60–75%), with a subsequent delay in the timing of snowmelt and increase in runoff response (Slater et al., 2001). The bias in the Phase 2(e) dataset with respect to the observations is much lower than that imposed in the Phase 2(d) experiment. Based on the findings of Phase 2(d), an underprediction of longwave radiation may result in overprediction of SWE with melt occurring too late. The ability of the models to predict SWE is discussed by Nijssen et al. (2003–this issue). It is expected that the sensitivity would be less in the colder arctic environment of the Torne–Kalix basin where mid-winter melt events are negligible, than in the warmer climate of Valdai.

4.4. Precipitation

4.4.1. Correction for gauge undercatch

Underprediction of the true quantity of solid precipitation by standard gauges in operation in high-latitude countries is a well-documented phenomenon,

as summarized by Sevruk and Hamon (1984). Systematic losses of solid precipitation due to aerodynamic, wetting and evaporation effects can reach 100% of the observed total precipitation, depending on gauge type, exposure and the time scale of interest (Goodison et al., 1998). Several methods have been proposed for correction of the Swedish manual gauges (Dahlström, 1973; Eriksson, 1983; Førland et al., 1996; Goodison et al., 1998; Alexandersson, 2001). Most of these methods require air temperature and wind speed, at a minimum. Because wind speed is only available at a few of the SMHI precipitation stations in the Torne–Kalix basin, precipitation catch correction for the Phase 2(e) experiment was based on constant monthly correction factors for each gauge developed by Alexandersson (2001). The corrections are designed to compensate for aerodynamic undercatch, wetting and evaporation losses. The coefficients developed by Alexandersson represent adjustments to the aerodynamic corrections of Eriksson (1983) and in general are somewhat smaller than the original corrections of Eriksson (1983), due in part to increased sheltering of the gauges over time. The mean monthly coefficients from Alexandersson (2001) were applied directly to the daily observations. Automatic stations (installed at seven of the stations beginning in 1995) were first adjusted by the monthly ratio of precipitation recorded at automatic stations to manual station observations for the overlapping period 1996–1999. Data from automatic stations subsequently were treated in the same way as observations from the manual stations. The distribution of precipitation stations in

Table 4

Differences in annual total gauge-corrected precipitation (1979–1998): [Førland et al. \(1996\)](#) and [Alexandersson \(2001\)](#)

Station	Minimum difference		Maximum difference		Mean difference		Mean annual gauge precipitation (mm)
	(mm)	(%)	(mm)	(%)	(mm)	(%)	
Haparanda	42.2	5.3	251.4	26.0	107.1	13.0	597
Katterjåkk	77.5	8.4	492.3	26.4	212.2	15.7	808
Naimakka	1.3	0.2	171.0	21.2	7.9	7.9	562

the mountainous areas of the study region is sparse and is unable to capture the variations of precipitation with elevation. Based on the recommendation of SMHI scientists, the catch corrected precipitation was lapsed to the grid cell elevation at a rate of 0.1%/m ([Graham, 2001](#)).

As part of the World Meteorological Organization (WMO) Solid Precipitation Measurement Intercomparison, the Finnish Meteorological Institute operated a site at the Jokioinen observatory in southern Finland ([Goodison et al., 1998](#)). [Førland et al. \(1996\)](#) developed gauge catch correction procedures for the Swedish standard manual gauge based on this experiment. The corrections for solid precipitation require air temperature and wind speed at the gauge orifice; correction of liquid precipitation requires wind speed, air temperature and precipitation intensity.

In order to provide an approximate assessment of uncertainties in the gage catch corrections, the corrections used in this experiment were compared to the procedure of [Førland et al. \(1996\)](#) for three manual stations located within the Torne–Kalix basin. For this comparison, the solid precipitation correction factor was calculated as an empirical function of observed air temperature and wind speed ([Førland et al., 1996](#)). Because hourly precipitation intensity is not observed at these stations, the aerodynamic correction factor for liquid precipitation was derived from a look-up table based on gauge exposure ([Førland et al., 1996](#)). An average of the correction factors was used for mixed precipitation (between 0 and 2 °C). Wetting losses were assigned a value of 0.02 mm per observation (once a day on precipitation days) and evaporation losses were based on monthly values in units of mm/day of precipitation ([Førland et al., 1996](#)).

Both the [Førland et al. \(1996\)](#) and the [Alexandersson \(2001\)](#) correction procedures were applied to the daily station data for the period 1979 through 1998.

The minimum, maximum and mean differences in annual totals for each station are summarized in [Table 4](#). The mean (1979–1998) monthly precipitation and corrections are shown in [Fig. 8](#). This comparison shows that the correction method of [Førland et al. \(1996\)](#) generally predicts more winter precipitation than that of [Alexandersson \(2001\)](#), with the magnitude of the differences depending on gauge exposure. It should be pointed out that the station where the largest discrepancies occur is in the mountains where measurement errors are most pronounced and appli-

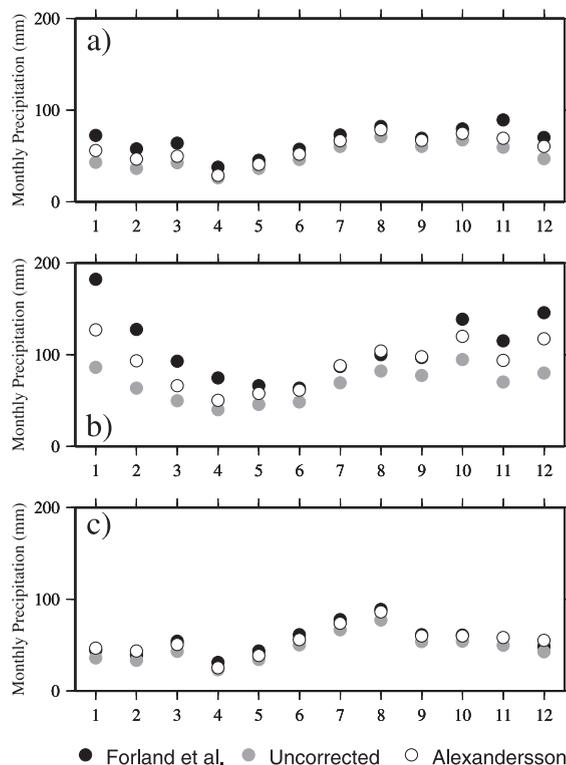


Fig. 8. Mean monthly total precipitation (1979–1998): (a) Haparanda station, (b) Katterjåkk station and (c) Naimakka station.

cation of standard correction procedures is difficult at best. The correction method of Førlund et al. (1996) is sensitive to the choice of the exposure coefficient for reducing wind speed to the height of the gauge orifice. For this comparison, we assigned values of exposure consistent with the original gauge classifications of Eriksson (1983). Alexandersson (2001) reduced the exposure classes of Eriksson based on local knowledge and updated location characteristics, because the corrections seemed too high in extreme cases. The corrections applied for this experiment rely on the best available data and there does not appear to be any obvious residual bias in the precipitation dataset.

4.4.2. Precipitation disaggregation

Sub-daily precipitation data for the Torne–Kalix basin are only available at a few automatic stations installed in the mid-1990s. For the model base runs, daily precipitation data were interpolated to hourly using a constant (1/24) apportionment. In some models, small quantities of constant precipitation may cause instabilities in interception evaporation. This is especially true if the precipitation occurs during periods of low cloud cover. To address these concerns, the organizers distributed a sub-daily precipitation data set for the rerun experiment (experiment 3) generated using a simplified statistical disaggregation technique. A cumulative distribution function (CDF) of event duration and start time, in 6-h increments, was created for each of the four automatic stations within the Torne–Kalix basin that recorded six-hourly precipitation. The derived CDF of the nearest station was used for each grid cell in the data set. For each 6-h bin with precipitation, the signal was further disaggregated to hourly values using a uniform probability distribution, while maintaining the daily totals. Based on observations at the four stations, 90% of the precipitation falls when the cloud cover exceeds 60%, so this method does not assign precipitation to hours with less than 60% cloud cover.

Due to the lack of hourly precipitation data in the basin, there is no way to verify the disaggregation procedure independently. Given that 40–60% of the annual precipitation falls as snow that is not immediately available for runoff, errors in the sub-daily precipitation signal are not expected to affect winter or spring seasonal or annual runoff production greatly. A similar experiment conducted for the Phase

2(c) intercomparison resulted in large percentage increases in the runoff ratio for some of the schemes. This was due in part to the small runoff ratio of the Arkansas–Red River basin. In most cases, changes in partitioning of water and energy were small (Wood et al., 1998). Because precipitation for much of the year is stored in the snowpack in the Torne–Kalix basin, the change in runoff ratio due to the disaggregation is expected to be considerably smaller than in the Arkansas–Red basin. Because precipitation is assigned to rain or snow based on the hourly air temperature, the precipitation disaggregation did result in a slight change in the annual snow fraction. On average for the period 1989 through 1998, the change in total precipitation due to round-off errors was approximately -0.05% . On average, 0.8% less precipitation was treated as snow using the precipitation disaggregation, as compared to the constant (1/24) apportionment. The maximum changes for individual years were an increase in snow of 0.5% in 1992 and a decrease of 2.7% in 1995.

5. Summary results

5.1. Calibration experiment

One of the primary motivations for an arctic model intercomparison was to provide information about the accuracy with which land-surface schemes can be used to estimate runoff from ungauged, high-latitude areas (WCRP, 1999a). The purpose of the calibration experiment was to test the extent to which calibration could improve the performance of the LSSs and the extent to which parameter transfer from such calibrated, gauged catchments can improve the estimation of runoff from similar, ungauged areas. The effect of calibration on model performance was explored for two sub-basins, the 566-km^2 Övre Abiskojokk and the 1341-km^2 Övre Lansjärv, shown in Fig. 2. In addition to the hourly forcing data, daily observed streamflow data were provided for 1989–1998 and modelers were instructed to calibrate their models to the extent possible with these limited data, following the same basic experiment design as described by Wood et al. (1998) for Phase 2(c). Only results using the rerun forcing data are described here. The catchments span multiple $1/4^\circ$ grid cells, so runoff from the contribu-

ting grid cells was summed and scaled according to the contributing area of each cell. Rather than routing the runoff produced by the various models from cell-to-cell, only monthly discharge was evaluated, which avoids the problem of accounting for channel travel times.

Twelve of the 21 models performed some sort of calibration for these basins, as summarized in Table 5. Eleven of these modeling groups provided the organizers with simulated streamflow both before and after calibration for Övre Abiskojokk and Övre Lansjärv. The nine groups that did not calibrate cited three primary reasons. First, there was the sense that changes to the parameters identified for calibration (mainly soil hydraulic properties) would not result in any real improvement in the results and it was therefore not worth the time expenditure. Second was the strategy that with only runoff as a calibration variable, improvement in runoff prediction might occur at the expense of other variables. Finally, some participants felt that it was desirable to test model performance in an operational sense, in which no calibration is feasible.

For the models that did calibrate, the agreement between the simulated and observed hydrographs, as

measured by the monthly root mean squared error (RMSE), was improved through calibration in almost all cases (Fig. 9a and b). In these figures, the vertical axis is the RMSE before calibration and the horizontal axis is the RMSE after calibration, so points above the dashed diagonal line represent an improvement in fit. Although the range in RMSE is approximately the same for both catchments, the effects of calibration were somewhat more consistent among models in the Övre Lansjärv catchment. The MATSIRO (model K) and SSiB (model R) models in particular showed similar substantial improvement in model performance following calibration in the Övre Abiskojokk catchment. The SSiB model calibration included the adjustment of the stability criterion to make the model less stable and an effective increase in the albedo of melting snow. By changing the energy partitioning, the SSiB calibration affected both the fit and the volume of simulated runoff. In contrast, calibration for MATSIRO focused on increasing the hydraulic conductivity and decreasing the saturated area fraction. These changes improved the RMSE of the simulation, with little change to the volume ratio. Changes in the simulated annual runoff volume were less pronounced for other models as well (Fig. 9c and

Table 5
Calibration techniques for models that participated in experiment 1

Model	Calibration technique	Transfer technique
CLASS	Increased saturated hydraulic conductivity by two orders of magnitude.	None
HY-SSiB	Changed from the Clapp and Hornberger (1978) soil properties, to the Cosby et al. (1984) parameters.	Used same parameters as calibration.
IHAS	Effective soil depth altered to match annual water balance and hydrograph shape.	None
ISBA	Included a sub-grid drainage scheme with default values of the drainage parameter.	Sub-grid drainage enacted with default parameter.
MATSIRO	Reduced exponential reduction of K_{sat} , enhanced surface K_{sat} , reduced TOPMODEL saturation area contribution factor (reduces runoff from saturated areas).	Used same parameters as calibration.
MECMWF	Depth for calculating relative soil saturation adjusted to optimize depth and phasing of peak discharge.	Same soil depth applied everywhere.
NSIPP	Adjusted exponential decay factor of saturated hydraulic conductivity.	Same decay factor applied everywhere.
SAST	Automatic calibration of all model parameters.	Used same parameters as calibration.
SEWAB	Applied a logarithmic profile for saturated conductivity.	Transferred without any changes.
SSiB	Melting rate was reduced; albedo increased for Abiskojokk; neglected stability correction for Lansjarv	Transferred parameters according to vegetation types in Lansjarv and Abiskojokk.
VIC	Adjusted surface storage drainage parameters and saturated hydraulic conductivity	Used same parameters as calibration.
VISA	Parameterized the exponential decay factor of saturated hydraulic conductivity as a function of elevation	Used a parameter between the values found for the two calibration catchments.

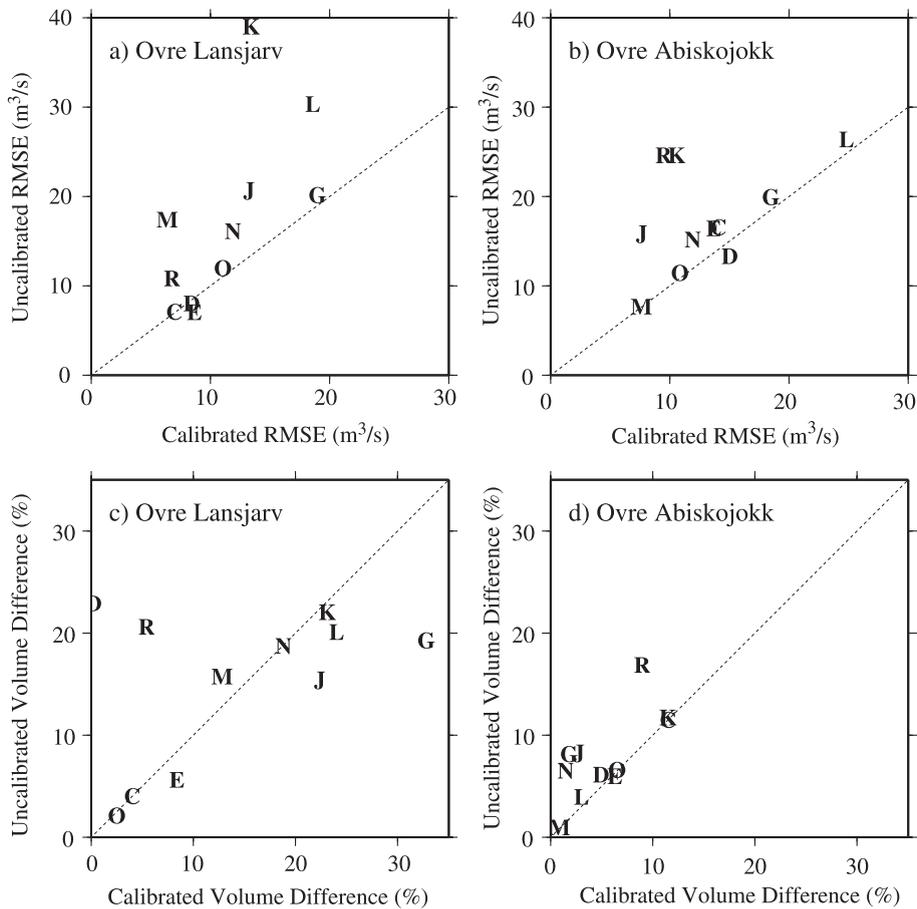


Fig. 9. Calibration basins runoff statistics: (a and b) uncalibrated and calibrated RMSE for Övre Lansjärv and Övre Abiskojokk, respectively, and (c and d) uncalibrated and calibrated percent difference in volume for Övre Lansjärv and Övre Abiskojokk, respectively. In this figure, values above the line represent a decrease in the absolute value of the percent difference between simulated and observed annual runoff volume with calibration. See Table 7 for model key.

d). In general, changes in model soil parameters tended to change the timing of runoff, but had little effect on the annual runoff volumes, particularly in Övre Lansjärv. One explanation for this lack of sensitivity of the annual runoff ratio to model parameters may be that evapotranspiration is primarily energy-limited in this environment. The experiment is somewhat non-conclusive, however, because some modeling groups may have emphasized hydrograph fit rather than volume during calibration.

Following calibration, modelers were asked to transfer their calibrated parameters to two validation catchments in order to test the potential improvement to be gained in simulation of ungauged basins. The

validation catchments were the 2622-km² Pello catchment and the 1472-km² Kaalasjärvi catchment (Fig. 2). Ten of the 21 LSSs used some calibration knowledge to determine the parameters of the validation basins, as summarized in Table 5. All 10 provided the organizers with simulated streamflow both before and after calibration for the Pello and Kaalasjärvi catchments. The methods of transfer varied, but most groups applied their changes globally without regard for soil or vegetation type. The remaining models simulated runoff with no changes to their default parameters. When compared to their results before parameter transfer, most models improved their representation of simulated runoff by transferring parameters from the vali-

validation catchments (Fig. 10a and b). Fig. 10c shows the results for all models together. Those models that did transfer parameters (indicated by capital letters) tend to perform better than those that did not transfer parameters (lower case letters), both with respect to RMSE and annual runoff volume. Prior to simulating the entire 218 grid cells, the 10 models that performed the parameter transfer experiment (indicated in Table 1) applied the same rationale to transfer the parameters and/or technique to the basin as a whole. As reported by Nijssen et al. (2003-this issue), the mean RMSE for the Torne and Kalix basins for those models that did not transfer parameters is approximately double the mean for models that did transfer parameters (1.54 and 0.77, respectively, for the Torne and 1.61 and 0.88 for the Kalix). The mean absolute bias is also lower for

those models that calibrated and transferred parameters. The improvements in RMSE seem to be influenced in large part by a better representation of low flows for those models that calibrated (Nijssen et al., 2003-this issue).

The calibration experiment indicates that calibration generally improves the fit of simulated hydrographs in these high-latitude catchments, and the volume to a lesser extent. The models have limited ability to correctly simulate runoff volumes for the two small catchments in the boreal forest (Pello and Övre Lansjarv), whether or not streamflow was provided or calibration attempted. For the small basins above the tree line, and for the larger basins as a whole, calibration (Övre Abiskojokk) and parameter transfer (Kaalasjarvi) offer potential for improving

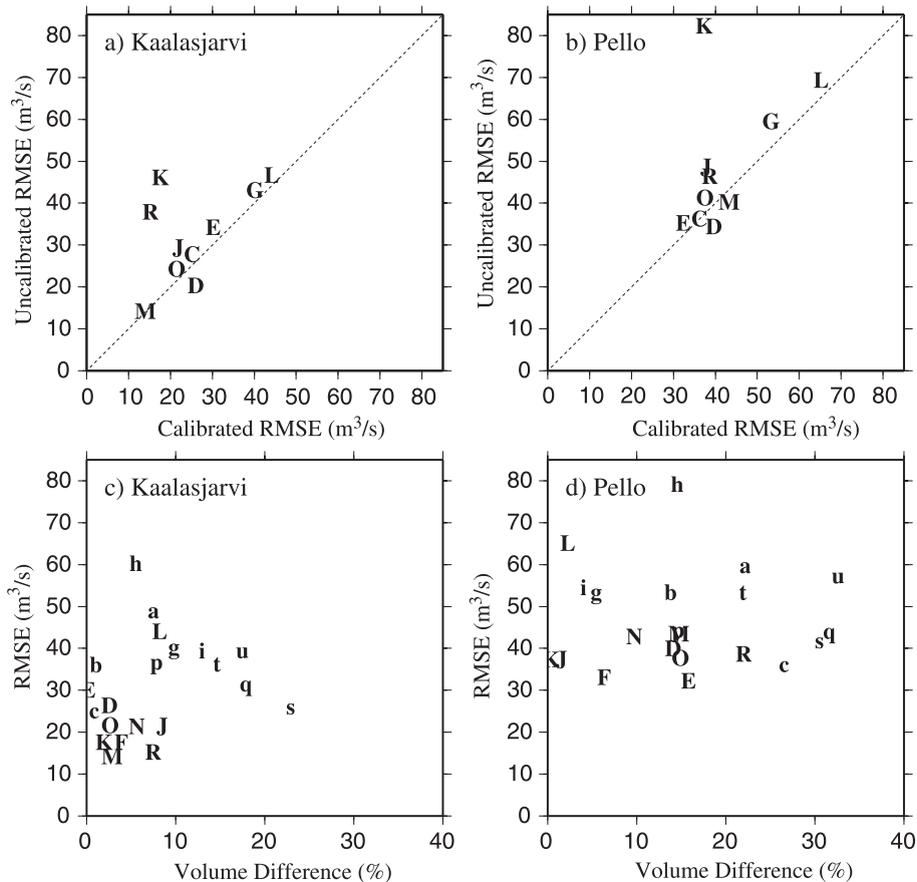


Fig. 10. Validation basin runoff statistics. (a and b) Uncalibrated and calibrated RMSE for Kaalasjärvi and Pello, respectively, for the models that calibrated and (c and d) final percent difference in runoff volume versus RMSE for Kaalasjärvi and Pello, respectively, for all models. The models, which are calibrated, are indicated by uppercase letters.

estimates of runoff timing, and possibly volume, from gauged and ungauged areas.

5.2. Energy balance components

The surface energy balance was calculated for all 21 models for all grid cells in the model domain, according to the following equation:

$$\varepsilon_{\text{eb}} = \text{SW}_{\text{net}} + \text{LW}_{\text{net}} - Q_{\text{le}} - Q_{\text{h}} - Q_{\text{g}} - \Delta\text{SoilHeat} \quad (3)$$

where SW_{net} , LW_{net} , Q_{le} , Q_{h} , and Q_{g} are the basinwide 10-year mean shortwave and longwave radiation, latent, sensible and ground heat fluxes, respectively. $\Delta\text{SoilHeat}$ represents the change in heat storage in the soil layer above which the ground heat flux is resolved. This is not a real storage term, but rather an explicit component of the ground heat flux for models that do

not resolve the ground heat flux at the soil/atmosphere interface. The ε_{eb} term represents the magnitude of non-closure of the energy balance for any time period. Water and energy balance closure over the period 1989–1998 was verified for each model. One model (IHAS) has an average annual energy balance error in the rerun experiment in excess of 3 W/m^2 , which is the consistency criterion used in earlier PILPS experiments (Fig. 11).

A significant portion of the Torne–Kalix basin lies above the Arctic Circle and the entire basin is consequently characterized by low available energy. Mean annual incident shortwave radiation ranges between 85 and 105 W/m^2 , while mean annual incident longwave radiation ranges between 265 and 280 W/m^2 . The presence of snow during a large part of the year, roughly from October until June, and the associated high albedo, resulted in a simulated mean annual net shortwave radiation averaged over the basin ranging from 54 to 73 W/m^2 . Simulated mean

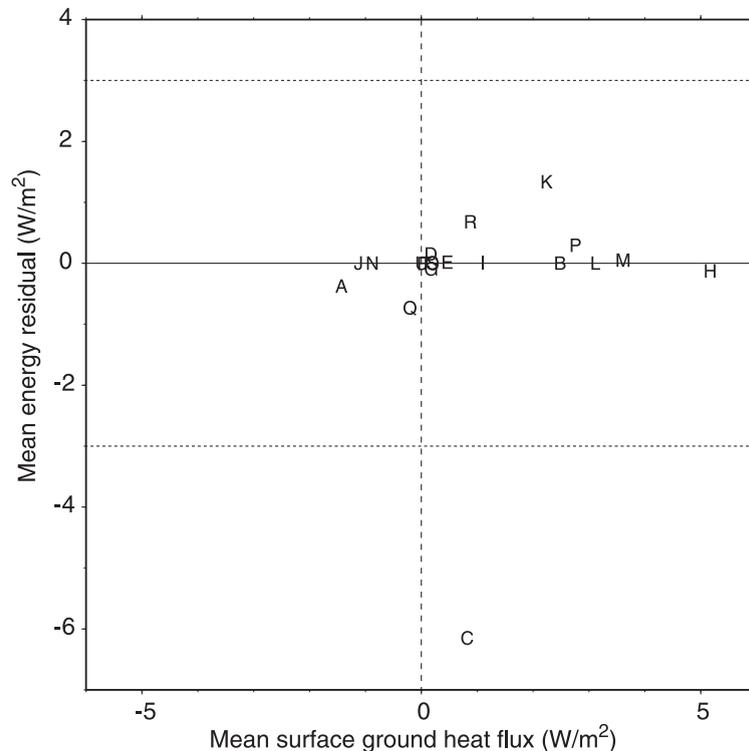


Fig. 11. Mean annual energy balance residuals (1989–1998) versus mean ground heat flux at the surface. See Table 7 for model legend. The dashed horizontal lines represent the 3 W/m^2 consistency criterion established by earlier PILPS experiments.

annual net longwave radiation also varied considerably between the models, ranging from -27 to -46 W/m^2 . Differences among models in surface temperature, and consequently net longwave radiation, were most marked during the winter months and were much smaller during the summer months when snow was absent (see, e.g., Nijssen et al., 2003-this issue).

As expected, net radiation was small for all models, with a range of 16 – 40 W/m^2 , as shown by the diagonal lines in Fig. 12. The mean annual simulated latent heat flux averaged over the basin ranged from 18 to 36 W/m^2 in the rerun. To clarify subsequent discussion, the models are categorized into four groups based on net radiation, as summarized in Table 6. Groups 1 and 4 have higher latent heat on average, and tend to over predict annual latent heat compared to that estimated from the annual water balance (indicated by the dashed vertical line). With the exception of model H (IBIS) models in Groups 2

Table 6

Summary of model groupings

Category	Net Radiation	Sensible Heat
Group 1	$R_{\text{net}} > 31$ W/m^2	$Q_{\text{h}} > 6.5$ W/m^2
Group 2	$22 < R_{\text{net}} < 27$ W/m^2	$-1.5 < Q_{\text{h}} < 6.5$ W/m^2
Group 3	$17 < R_{\text{net}} < 22$ W/m^2	$-7.5 < Q_{\text{h}} < 0.5$ W/m^2
Group 4	$16 < R_{\text{net}} < 26$ W/m^2	$Q_{\text{h}} < -7.5$ W/m^2

and 3 all simulate average annual latent heat within 5 W/m^2 of the estimate from the basin water balance. Group 3 has lower net radiation on average than Group 2. The latent heat flux peaked for all models during the summer months, but significant differences existed among the models during all months.

The simulated mean annual sensible heat is for the most part small, but is positive for some models and negative for others, with a fairly equal distribution around zero (Fig. 12). Mean annual sensible heat

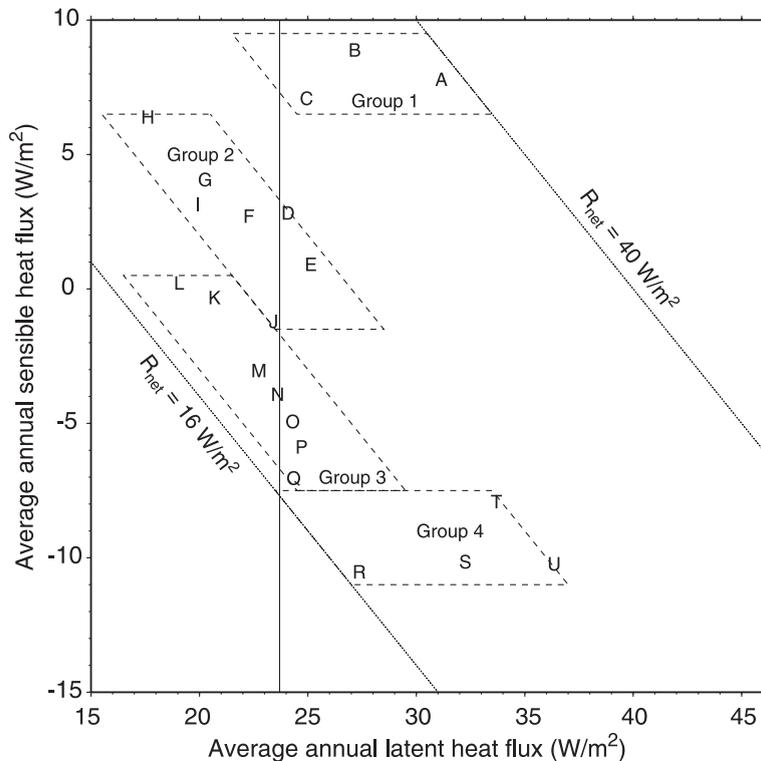


Fig. 12. Annual average latent versus sensible heat. Dotted diagonal lines represent lines of equal net radiation, assuming, on average, that ground heat flux and storage changes are zero. The solid vertical line is the annual average latent heat flux estimated from the basin water balance. The groupings were established for discussion purposes based on annual net radiation and sensible heat.

fluxes in the rerun ranged from -11 to $+9$ W/m^2 . A negative sensible heat flux implies that the atmosphere acted as a source of energy for evapotranspiration and sublimation. In off-line simulations such as this, the near-surface temperature cannot change in response to this simulated energy sink, which may have important implications for the resulting water and energy balance.

For most models, the annual latent heat flux is greatest in the more southerly, lower elevations, rather than the high elevations that receive greater precipitation (as shown in Fig. 13). This suggests that on an annual basis latent heat (evapotranspiration) is limited by available energy, at least in the northern part of the basin.

5.3. Water balance components

Errors in the basin average water balance for each model were calculated according to:

$$\varepsilon_{\text{wb}} = P - E - Q_s - Q_{\text{sb}} - \Delta\text{Storage} \quad (4)$$

where P is the sum of snow and rain precipitation, E is evapotranspiration from all sources and Q_s and Q_{sb} are surface and subsurface runoff, respectively. The change in storage terms includes the total change in snow water equivalent (SWE), soil moisture, surface water storage, and canopy interception storage. One model (RCA) exceeds the 3-mm consistency criterion used in earlier PILPS experiments, as shown in Fig. 14.

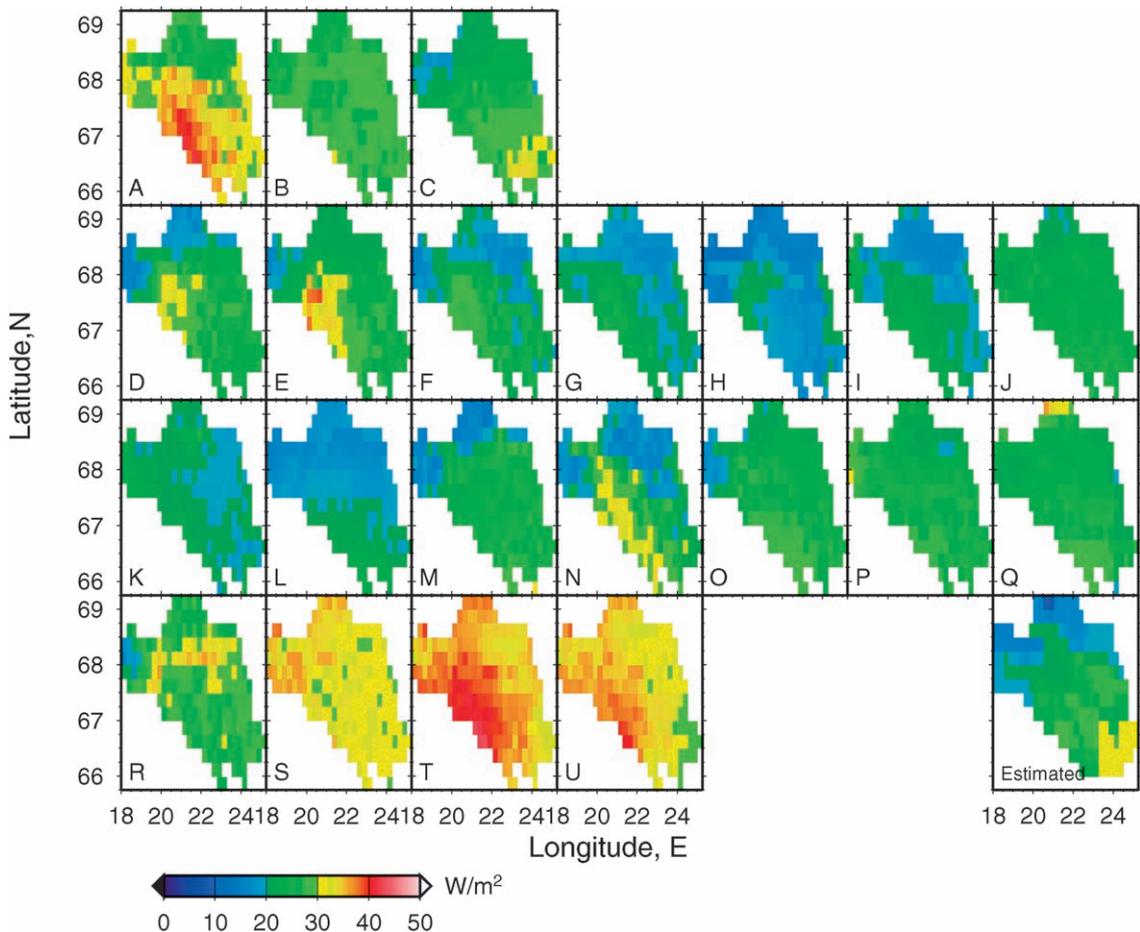


Fig. 13. Mean annual latent heat flux for the model reruns. The last panel shows annual latent heat flux estimated from basin water balance.

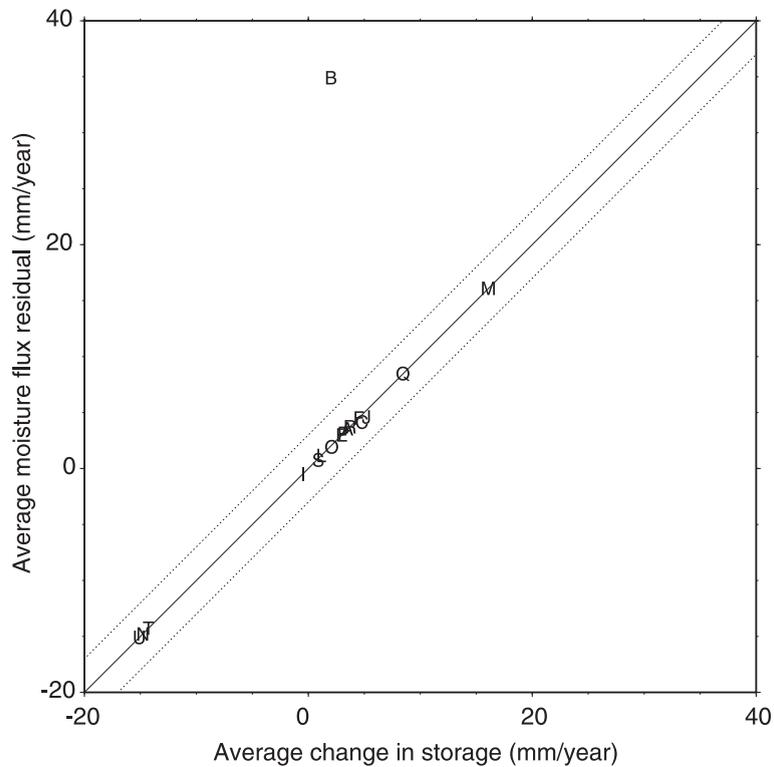


Fig. 14. Mean moisture flux residuals ($P - E - Q_s - Q_{sb}$, see Eq. (4)) versus the total change in storage (1989–1998). The dashed lines represent the limit of the 3-mm consistency criterion for water balance closure established by previous PILPS experiments.

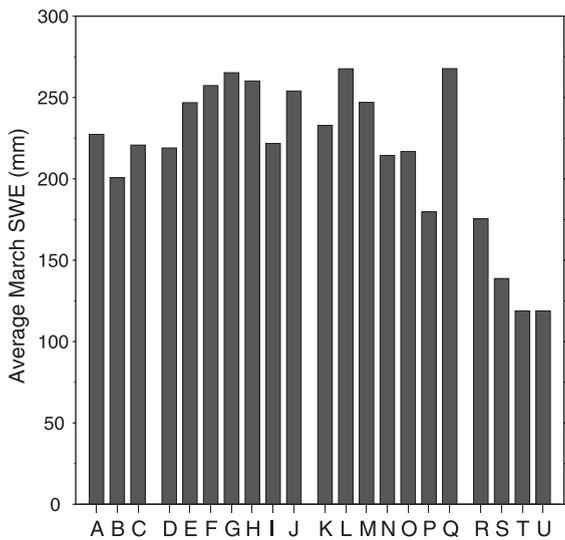


Fig. 15. Basin average snow water equivalent (SWE) for March (surrogate for maximum) for the period (1989–1998).

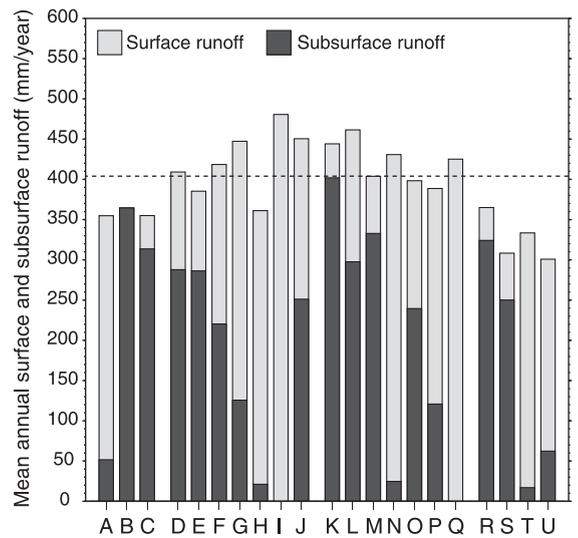


Fig. 16. Total basin mean annual surface and subsurface runoff. The dashed horizontal line represents mean annual runoff at the mouths of the Torne and Kalix Rivers combined.

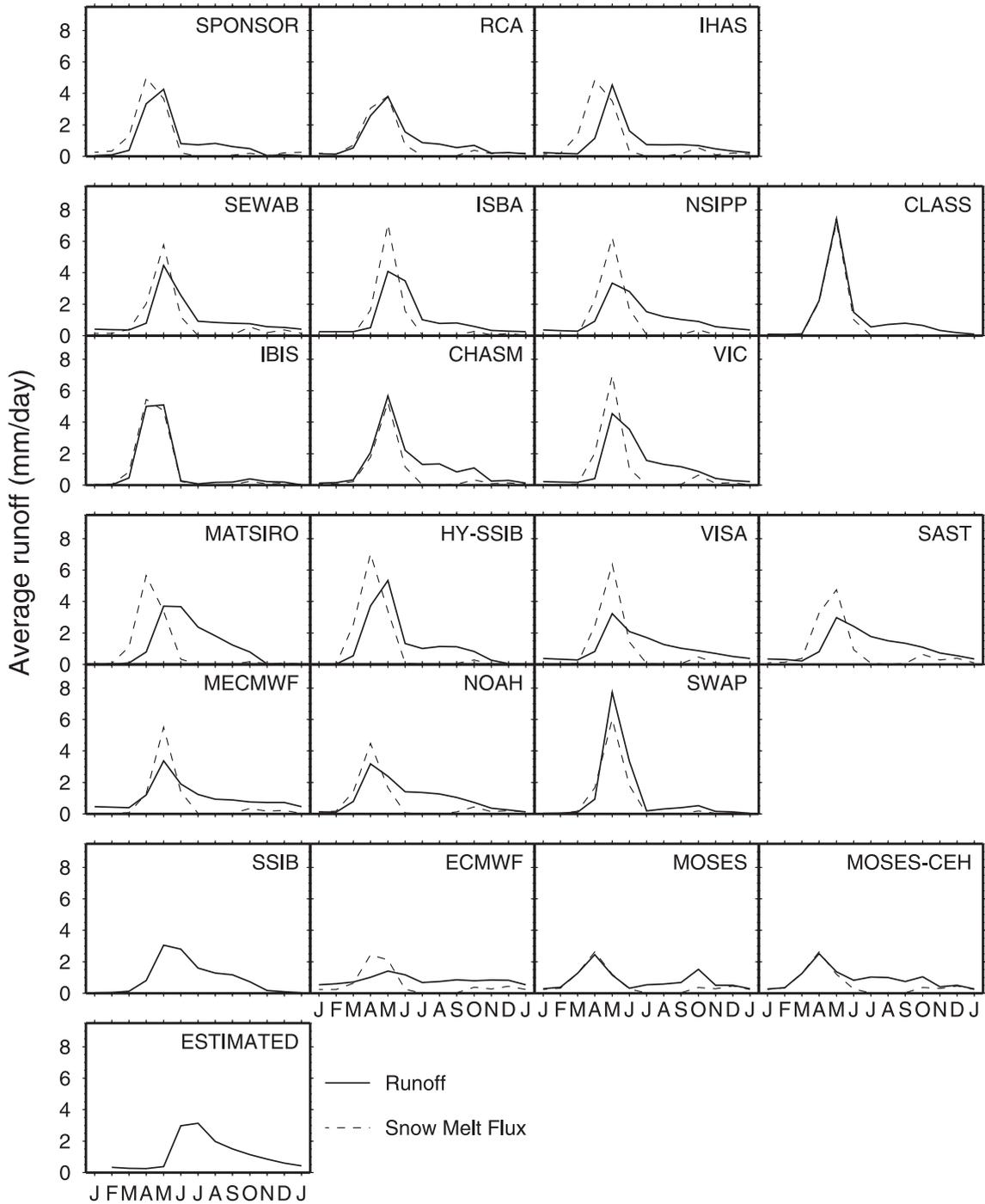


Fig. 17. Total mean monthly runoff (not routed streamflow) and mean monthly snow melt flux.

Perhaps most striking is the large variation in predicted SWE among models, and the effect of these differences on the modeled energy and moisture fluxes throughout the year. Averaged over the 10 years, mean March simulated SWE ranged from 119 to 268 mm, with a mean of 217 mm, as shown in Fig. 15. Nine of the models have maximum snow accumulation in March on average, the remaining 12 models peak in April. Mean March SWE is used as a surrogate for maximum accumulation because it is closer to the annual maximum for the majority of the models. Mid-winter melt events are infrequent over most of the basin and have relatively little effect on maximum snow accumulation. The models with high annual latent heat flux and negative sensible heat (Group 4) tend to have the lowest snow accumulation. These models also have larger annual net shortwave radiation, on average (see Nijssen et al., 2003-this issue). However, due to the very low values of incident shortwave radiation during

the snow accumulation season (September–March), this cannot be the reason for the low snow accumulation. In contrast, models in Group 1 have similar annual latent heat as Group 4, but positive sensible heat and subsequently, less negative net longwave radiation. As a result, Group 1 models have larger annual net radiation than those in Group 4. These models do not have maximum SWE that differs markedly from Groups 2 and 3. The difference is due to the seasonal cycle of latent heat—Group 1 models have higher summer latent heat on average, whereas the models in Group 4 tend to have higher winter latent heat (Nijssen et al., 2003-this issue). For the MOSES model, Essery and Clark (2003-this issue) found that this is due in part to the treatment of snow intercepted by the vegetation canopy.

The volume and shape of the predicted seasonal runoff hydrographs varied considerably among models, with annual runoff ranging from 301 to 481 mm for the reruns, as shown in Fig. 16. The observed

Table 7
Summary of model structure

Model	Snow model	Stability correction	Frozen soils	Vegetation scheme	Surface water representation
<i>Group 1</i>					
(A) SPONSOR	Bulk layer	Gradient Richardson	Diffusion	Average	None
(B) RCA	Composite	Bulk Richardson	Diffusion	Average	None
(C) IHAS	Multi-layer	None	Temp index	Average	None
<i>Group 2</i>					
(D) SEWAB	Bulk layer	Bulk Richardson	Diffusion	Average	Evaporation
(E) ISBA	Multi-layer	Flux Richardson	Temp index	Average	None
(F) NSIPP	Multi-layer	Bulk Richardson	Diffusion	Average	None
(G) CLASS	Bulk layer	Bulk Richardson	Diffusion	Average	None
(H) IBIS	Multi-layer	Bulk Richardson	Diffusion	Average	None
(I) CHASM	Composite	Bulk Richardson	None	Mosaic	None
(J) VIC	Multi-layer	Bulk Richardson	Diffusion	Mosaic	Evaporation/Storage
<i>Group 3</i>					
(K) MATSIRO	Multi-layer	Monin-Obukhov	Diffusion	Majority	Evaporation/Storage
(L) HY-SSiB	Multi-layer	Bulk and Monin-Obukhov	Temp index	Majority	None
(M) VISA	Multi-layer	Monin-Obukhov	Diffusion	Mosaic	Evaporation
(N) SAST	Multi-layer	Bulk Richardson	Diffusion	Average	Evaporation
(O) MECMWF	Bulk layer	Bulk and Monin-Obukhov	Diffusion	Mosaic	None
(P) NOAA	Composite	Paulson (1970)	Diffusion	Dominant	None
(Q) SWAP	Bulk layer	Monin-Obukhov	Other	Partial average	None
<i>Group 4</i>					
(R) SSiB	Multi-layer	Bulk Richardson	Diffusion	Mosaic	None
(S) ECMWF	Bulk layer	Bulk and Monin-Obukhov	Diffusion	Mosaic	None
(T) MOSES-CEH	Composite	Bulk Richardson	Diffusion	Mosaic	Evaporation
(U) MOSES	Composite	Bulk Richardson	Diffusion	Mosaic	Evaporation

mean annual discharge of the Torne–Kalix system is 403 mm. Snowmelt is the single largest factor controlling annual runoff production in the Torne–Kalix basin. Not surprisingly, therefore, the models with high SWE tend to have the highest annual runoff (Groups 2 and 3). Group 1 does tend to have lower annual runoff despite near-average snow accumulation, due to higher than average summer evaporation. However, one of the three Group 1 models does not close the water balance. There is no apparent relationship between annual runoff and the ratio of surface to subsurface flow; all four groups include models dominated by either surface or subsurface runoff (Fig. 16). This suggests that model representation of subsurface runoff does not, in general, control simulated annual runoff volume in the Torne–Kalix basin. This may be due either to the inability of models to represent true groundwater flow, or energy-limitations on evapotranspiration, or both.

On the other hand, models with similar snow accumulations simulated different mean monthly runoff due to differences in the treatment of surface and subsurface storage, as shown in Fig. 17. Two of the models (VIC, J, and MATSIRO, K) explicitly simulated surface storage in lakes (see Table 7), and while these models tended to simulate less moisture storage in the soil column, the resulting shape of the simulated hydrographs for these models did not differ much from those without lakes. Models with a high proportion of surface flow, as expected, have a fast runoff response on average (e.g., models IBIS, CHASM, SWAP, MOSES and MOSES-CEH). Although the runoff peak for these models is not always higher than those predominated by subsurface runoff (e.g., models IHAS, SEWAB, ISBA, NSIPP, VIC, MATSIRO, HY-SSiB), it is characterized by a shape that is a direct translation of the snowmelt signal. Group 4 models tend to underestimate runoff due to the very low snow accumulation.

6. Discussion: model representation of basin hydrology

6.1. Snow regime

Snow cover accumulation, redistribution and ablation dominate the land-surface hydrology of high

latitudes in general, and the Torne–Kalix basin in particular. Although rain accounts for between 40% and 60% of the annual precipitation in this basin, snow is proportionately more important to runoff due to low rates of winter evapotranspiration and the timing of the melt water release.

Sublimation during blowing snow events in regions of arctic tundra and prairies has been shown to affect a substantial percentage of annual snow precipitation (Woo, 1982; Kane et al., 1991; Hinzman et al., 1996). Snow redistribution by wind is also significant in the alpine zone above tree line (e.g., Greene et al., 1999), and we anticipate that sublimation during blowing snow events may be important above the tree line in the Torne–Kalix basin. This alpine region is further characterized by an increase in snow accumulation with elevation due to orographic enhancement of precipitation and decreasing temperature. It has been observed in the North American Rocky Mountains that the resulting gradient of snow depth versus elevation continues to steepen throughout the snow season as snow continues to accumulate at the highest elevations while differential melting occurs in the valleys (Storr and Golding, 1974). Because blowing snow is not represented by any of the models, we would expect a tendency toward overprediction of snow cover, especially above the tree line.

In addition, boreal forest canopies may intercept up to 60% of snowfall (Pomeroy and Gray, 1995). Increased exposure and turbulent transfer in the canopy relative to the ground snowpack can result in sublimation losses from intercepted snow of between 30% and 40% of annual snowfall (Krestovskiy et al., 1972; Pomeroy and Schmidt, 1993).

In PILPS Phase 2(d), Slater et al. (2001) found that a significant portion of the scatter in maximum snow accumulation was due to the models' response to mid-season ablation events. Such events do not often occur in the Torne–Kalix basin, which makes differences in sublimation proportionately more important in explaining between-model differences in snow accumulation. Fig. 13 indicates that large among-model differences in annual evapotranspiration occur both above and below the tree line (see Fig. 2). This suggests that the sensitivity of models to the aerodynamic resistance and atmospheric stability corrections in areas of no overstorey are at least as important as the

sensitivity to representation of canopy interception and heat fluxes.

6.2. Frozen soil and surface storage

The Torne–Kalix drainage basin is underlain primarily by seasonally frozen ground and some local permafrost at the highest altitudes. Percolation of snow melt water may raise soil surface temperatures to 0 °C, but soils usually remain frozen until the snow cover is depleted (Woo and Steer, 1983; Woo and Winter, 1993). The presence of frozen water limits infiltration into the soil and decreases the melt water storage capacity of the soils. Therefore, an increase in the proportion of surface flow is typical of both permafrost and seasonally frozen regimes during snowmelt.

Numerous lowland lakes and bogs are also characteristic of the boreal forest vegetation of these basins. According to Church (1974), northern lakes and wetlands create a muskeg (peat bog) regime, in which high

flows are attenuated by the absorption capacity of the muskeg vegetation and resistance to flow by surface conditions. The storage capacity of wetlands, however, is often limited by frozen ground, because the high ice content wetland soils are slow to thaw (Roulet and Woo, 1986). Therefore, flow attenuation by wetlands is usually highest for summer rainfall after the peat is thawed (Roulet and Woo, 1986; Woo, 1988).

The numerous lakes and peat bogs in the Torne–Kalix basin appear to exert a strong influence on observed hydrographs even at the highest elevations, as shown in Fig. 18. For the headwater catchments above the tree line, the hydrographs become more attenuated with increasing open water areas. Below the tree line, the effect is convoluted by the attenuating effects of basin size, but still appears evident. Fig. 18 also indicates that the runoff volume normalized by basin area is not dependent on the open water percentage. There is a distinct difference in runoff rate between the headwaters and the lower catchments, but

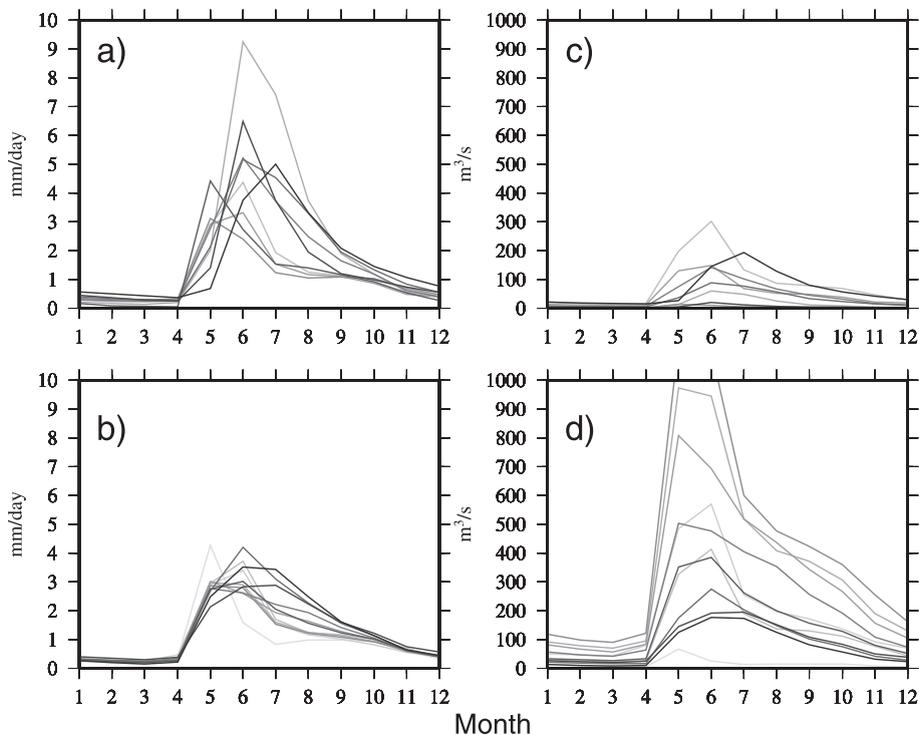


Fig. 18. Observed monthly hydrographs for 20 catchments in the Torne–Kalix basin: (a) normalized (mm/day) above treeline and (b) above treeline (m^3/s); (c) normalized (mm/day) below treeline and (d) below treeline (m^3/s). Line color darkens with increasing percent open water area in each catchment from 2% to 15.4%.

no clear differentiation with lake area. This is in agreement with the results of the calibration experiment that suggest that holding water longer either as surface storage or soil storage appears to not greatly increase total evapotranspiration.

7. Conclusions

This paper describes the structure of the Phase 2(e) model intercomparison, conducted over the Torne–Kalix Rivers in northern Scandinavia. As is often the case in high-latitude studies, the experiment is limited, in part, by the quality of atmospheric forcing data available. Known problems with the meteorologic data for the Torne–Kalix basin include the following:

- Wind-induced errors in the observation of solid precipitation resulted in the under-observation of solid precipitation that had to be corrected in the gridded dataset.
- Sub-daily precipitation data are scarce and virtually non-existent for the historic record, so precipitation was disaggregated statistically using the limited available data.
- Incoming radiation components are not regularly observed and must be estimated. An inappropriate estimation technique resulted in a large bias in the original forcing data that had to be replaced after the experiment was underway.

An analysis of the forcing dataset in conjunction with other available data products indicates that there are no significant biases in the gridded data. It is therefore appropriate to analyze the ability of the participating LSSs to represent this high-latitude environment.

Analysis of the Phase 2(e) model results, in the context of Phase 2(d) results and the hydrology of the Torne–Kalix basin, leads us to a set of working hypotheses regarding the functioning of LSSs in high-latitude regimes.

- At high latitudes, available net radiation rather than available water limits latent heat. The majority of the LSSs are able to capture this characteristic.
- Even though the largest evaporation rates occur in the summer (June, July and August), inter-model

variations in mean annual runoff are primarily related to the model-predicted winter snow sublimation. Models with higher rates of sublimation (hence, lower accumulations of snow on average) predict lower annual runoff.

- The formulation of aerodynamic resistance and stability corrections in areas of no overstory are at least as important as the sensitivity to representation of canopy interception in explaining intermodel differences in winter evaporation.
- Storage of snowmelt runoff either in the snowpack, soil column or surface ponding primarily influences the timing of the peak rates of both runoff and turbulent fluxes, rather than the absolute magnitude. There is no direct relationship between the ratio of subsurface to surface flow and the volume of annual runoff.
- Simulation of small high-latitude catchments dominated by peatlands and bogs by LSSs continues to be somewhat problematic. However, the LSSs show greater potential for estimation of runoff from ungauged regions with no overstory and limited soil overburden, which are perhaps more typical of much of the ungauged portion of the Arctic drainage basin.

These hypotheses are further explored through further analysis of the simulated results with respect to available observations throughout the Torne–Kalix basin by [Nijssen et al. \(2003-this issue\)](#). [Bowling et al. \(2003-this issue\)](#) relate the observed model performance to properties of the LSSs through the use of a simple bucket-model characterization of each LSS.

Acknowledgements

This paper was improved by the careful review of Ann Henderson-Sellers and one anonymous reviewer. Many of the authors were assisted in their modeling efforts by more contributors than can be listed here. Hans Alexandersson of the Swedish Meteorological and Hydrological Institute provided critical expertise with regard to catch correction of Swedish precipitation records. Jan Polcher provided support for the first use of the ALMA data exchange conventions. Travel support for the Phase 2(e) workshop held in Seattle, WA, was provided in part by the World Climate Research Program (WCRP) and by the Joint Institute

for the Study of the Atmosphere and Ocean (JISAO) at the University of Washington funded under NOAA Cooperative Agreement NA17RJ1232, as part of the GEWEX Continental-scale International Project (GCIP).

References

- Aagaard, K., Carmack, E.C., 1989. The role of sea ice and other freshwater in the Arctic circulation. *J. Geophys. Res.* 94c, 14485–14498.
- Alexandersson, H., 2001. Swedish Meteorological and Hydrological Institute, personal communication.
- Batjes, N.H., 1995. A Homogenized Soil Data File for Global Environmental Research: A Subset of FAO, ISRIC and NRCS Profiles International Soil Reference and Information Center (ISRIC), Wageningen, The Netherlands.
- Boone, A., Etchevers, P., 2001. An inter-comparison of three snow schemes of varying complexity coupled to the same land surface and macro-scale hydrologic models. *J. Hydrometeorol.* 2 (4), 374–394.
- Boone, A., Masson, V., Meyers, T., Noilhan, J., 2000. The influence of the inclusion of soil freezing on simulation by a soil–atmosphere-transfer scheme. *J. Appl. Meteorol.* 9, 1544–1569.
- Bowling, L.C., Nijssen, B., Lettenmaier, D.P., Polcher, J., Koster, R.D., Lohmann, D., 2003. Simulation of high-latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e): 3. Sensitivity analysis. *Glob. Planet. Change* 38, 55–71.
- Broecker, W.S., 1997. Thermohaline circulation, the Achilles heel of our climate system: will man-made CO₂ upset the current balance? *Science* 278, 1582–1588.
- Brutsaert, W., 1975. On a derivable formula for long-wave radiation from clear skies. *Water Resour. Res.* 11, 742–744.
- Carlsson, B., 1999. Some facts about the Torne and Kalix River Basins, Swedish Meteorology and Hydrology Institute, Hydrology Report No. 80.
- Carter, A.J., Scholes, R.J., 1999. Generating a global database of soil properties, Environmentek CSIR, South Africa.
- Cherkauer, K.A., Lettenmaier, D.P., 1999. Hydrologic effects of frozen soils in the Upper Mississippi River basin. *J. Geophys. Res.* 104d, 19599–19610.
- Church, M., 1974. Hydrology and permafrost with reference to northern North America. Proceedings Workshop Seminar on Permafrost Hydrology, Can. Nat. Comm., IHD, Ottawa, pp. 7–20.
- Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. *Water Resour. Res.* 14 (4), 601–604.
- Cosby, B.J., Hornberger, G.M., Clapp, R.B., Ginn, T.R., 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* 20 (6), 682–690.
- Dahlström, B., 1973. Investigation of errors in rainfall observations. Dept. of Meteorol., Univ. Uppsala, report No. 34, 31 pp.
- Eagleson, P.S., 1970. *Dynamic Hydrology* McGraw-Hill, New York.
- Eriksson, B., 1983. Data rörande Sveriges Nederbörds-klimat- Normalvärden för perioden 1951–80 (Data concerning the precipitation climate of Sweden-Normal values for the period 1951–80, in Swedish). Rapport 1983:28, Swedish Meteorological and Hydrological Institute, Norrköping, 92 pp.
- Essery, R., Clark, D.B., 2003. Developments in the MOSES land-surface model for PILPS 2e. *Glob. Planet. Change* 38, 161–164.
- FAO, 1995. The Digital Soil Map of the World, Version 3.5 Food and Agricultural Organization, Rome, Italy.
- Förland, E.J., Allerup, P., Dahlström, B., Elomaa, E., Jónsson, T., Madesn, H., Perälä, J., Rissanen, P., Vedin, H., Vejen, F., (Eds.), 1996. Manual for operation correction of Nordic precipitation, Norwegian Meteorological Institute, Report NR. 24/96.
- Freden, C., 1994. Geology. In: Freden, C. (Ed.), National Atlas of Sweden. Royal Swedish Academy of Sciences, Stockholm, pp. 106–119, 134–135 and 143–149.
- Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L., Fu, C., 2001. Regional climate information—evaluation and projects. In: Houghton, J.D., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., Johnson, C.A. (Eds.), *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge Univ. Press, Cambridge, UK. 881 pp.
- Goodison, B.E., Louie, P.Y.T., Yang, D., 1998. WMO solid precipitation measurement intercomparison, Final Report, WMO Instrument and Observing Methods Report No. 67, 212 pp.
- Graham, L.P., 2001. Swedish Meteorological and Hydrological Institute, personal communication.
- Greene, E.A., Liston, G.E., Pielke, R.A., 1999. Simulation of above treeline snowdrift formation using a numerical snow-transport model. *Cold Reg. Sci. Technol.* 30, 135–144.
- Hansen, M.C., DeFries, R.S., Toenshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* 21, 1331–1364.
- Henderson-Sellers, A., Yang, Z.L., Dickinson, R.E., 1993. The project for intercomparison of land-surface parameterization schemes. *Bull. Am. Meteorol. Soc.* 74, 1335–1349.
- Henderson-Sellers, A., Pitman, A.J., Love, P.K., Irannejad, P., Chen, T.H., 1995. The project for intercomparison of land-surface parameterization schemes (PILPS): phases 2 and 3. *Bull. Am. Meteorol. Soc.* 76, 489–503.
- Hinzman, L.D., Kane, D.L., Benson, C.S., Everett, K.R., 1996. Energy balance and hydrological processes in an Arctic watershed. *Ecol. Stud.* 120, 131–154.
- IAPO, 1998. Report on the ACSYS Workshop on status and directions for the Arctic runoff data base, Koblenz, Germany, 26–27 February 1998, IAPO Informal report number 1/1998.
- Idso, S.B., 1981. A set of equations for full spectrum and 8–14-m and 10.5–12.5-m thermal radiation from cloudless skies. *Water Resour. Res.* 17 (1), 295–304.
- Jin, J., Gao, X., Sorooshian, S., Yang, Z.-L., Bales, R.C., Dickinson, R.E., Sun, S.-F., Wu, G.-X., 1999a. One-dimensional snow water and energy balance model for vegetated surfaces. *Hydrol. Process.* 13 (14–15), 2467–2482.
- Jin, J., Gao, X., Yang, Z.-L., Bales, R.C., Sorooshian, S., Dick-

- inson, R.E., Sun, S.-F., Wu, G.-X., 1999b. Comparative analyses of physically-based snowmelt models for climate simulations. *J. Clim.* 12, 2643–2657.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woolen, J., Zhu, Y., Leetma, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* 77, 437–471.
- Kane, D.L., Hinzman, L.D., Benson, C.S., Liston, G.E., 1991. Snow hydrology of a headwater Arctic basin, I. *Water Resour. Res.* 27, 1099–1109.
- Krestovskiy, O.I., Postnikov, A.N., Sergeeva, A.G., 1972. Evaporation during the snow melting and flood period in spring. *Sov. Hydrol.* 5, 439–451.
- Lynch, A., McGinnis, D., Bailey, D., 1998. Snow-albedo feedback and the spring transition in a regional climate system model: influence of land surface model. *J. Geophys. Res.* 103 (D22), 29037–29049.
- Mitchell, K., Houser, P., Wood, E., Schaake, J., Tarpley, D., Lettenmaier, D., Higgins, W., Marshall, C., Lohmann, D., Ek, M., Cosgrove, B., Entin, J., Duan, Q., Pinker, R., Robock, A., Habets, F., Vinnikov, K., 1999. GCIP land data assimilation system (LDAS) project now underway. *GEWEX News* 9 (4), 3–6.
- Mocko, D.M., Sud, Y.C., 2001. Refinements to SSiB with an emphasis on snow-physics: evaluation and validation using GSWP and Valdai data. *Earth Interact.* 5 (5-001) (31 pp.).
- Mueller, L., 1998. Meteorological data on a horizontal grid, http://www.smhi.se/sgn0102/bhdc/metdata_3h_grid.htm.
- Myneni, R.B., Nemani, R.R., Running, S.W., 1997. Estimation of global leaf area index and absorbed PAR using radiative transfer models. *IEEE Trans. Geosci. Remote Sens.* 35 (6), 1380–1393.
- Nijssen, B., Schnur, R., Lettenmaier, D.P., 2001. Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980–1993. *J. Clim.* 14, 1790–1808.
- Nijssen, B., Bowling, L.C., Lettenmaier, D.P., Clark, D.B., El Maayar, M., Essery, R., Goers, S., Gusev, Y.M., Habets, F., van den Hurk, B., Jin, J., Kahan, D., Lohmann, D., Ma, X., Mahanama, S., Mocko, D., Nasonova, O., Niu, G., Samuelsson, P., Shmakin, A.B., Takata, K., Verseghy, D., Viterbo, P., Xia, Y., Xue, Y., Yang, Z., 2003. Simulation of high-latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e): 2. Comparison with observations. *Glob. Planet. Change* 38, 31–53.
- Niu, G.Y., Yang, Z.L., 2003. A physically based multi-layer snow parameterization for the NCAR CCM3/LSM: Part I. Validation and sensitivity tests in stand-alone experiments. *J. Clim.* (in review).
- Paulson, C.A., 1970. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteorol.* 9, 857–861.
- Polcher, J., Cox, P., Dirmeyer, P., Dolman, H., Gupta, H., Henderson-Sellers, A., Houser, P., Koster, R., Oki, T., Pitman, A., Viterbo, P., 2000. GLASS; global land–atmosphere system study. *GEWEX News* 10 (2), 3–5.
- Pomeroy, J.W., Gray, D.M., 1995. Snowcover accumulation, relocation and management, National Hydrology Research Institute Science Report No. 7, NHRI Environment Canada, Saskatoon, 144 pp.
- Pomeroy, J.W., Schmidt, R.J., 1993. The use of fractal geometry in modelling intercepted snow accumulation and sublimation. *Proc. Eastern Snow Conference*, vol. 50, pp. 1–10.
- Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of soil water properties. *Trans. ASAE* 25 (5), 1316–1320.
- Roulet, N.T., Woo, M.K., 1986. Hydrology of a wetland in the continuous permafrost region. *J. Hydrol.* 89 (1–2), 73–91.
- Samuelsson, P., Bringfelt, B., Graham, L.P., 2003. The role of aerodynamic roughness for the relationship between runoff and snow evaporation in land-surface schemes. *Glob. Planet. Change* 38, 93–99.
- Schlosser, C.A., Robock, A., Vinnikov, K.Ya., Speranskaya, N.A., Xue, Y., 1997. 18-year land surface hydrology model simulations for a midlatitude grassland catchment in Valdai, Russia. *Mon. Weather Rev.* 125, 3279–3296.
- Schlosser, C.A., Slater, A.G., Robock, A., Pitman, A.J., Vinnikov, K.Y., Henderson-Sellers, A., Speranskaya, N.A., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P., de Rosnay, P., Desborough, C.E., Dickenson, R.E., Dai, Y., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y., Habets, F., Kim, J., Koren, V., Kowalczyk, E., Nasonova, O., Noilhan, J., Schaake, J., Shmakin, A.B., Smirnova, T.G., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z., 2000. Standalone simulations of a boreal hydrology with land surface schemes used in atmospheric models: PILPS Phase 2(d). *Mon. Weather Rev.* 128, 301–321.
- Sellers, P.J., Meeson, B.W., Hall, F.G., Asrar, G., Murphy, R.E., Schiffer, R.A., Bretherton, F.P., Dickinson, R.E., Ellingson, R.G., Field, C.B., Huemmrich, K.F., Justice, C.O., Melack, J.M., Roulet, N.T., Schimel, D.S., Try, P.D., 1995. Remote sensing of the land surface for studies of global change: models–algorithms–experiments. *Remote Sens. Environ.* 51 (1), 3–26.
- Sevruk, B., Hamon, W.R., 1984. International comparison of national precipitation gauges with a reference pit gauge. *WMO Instrument and Observing Methods Report*, vol. 17. WMO, Geneva. 111 pp.
- Shepard, D.S., 1984. Computer mapping: the SYMAP interpolation algorithm. In: Gaile, G.L., Wilmott, C.J. (Eds.), *Spatial Statistics and Models*. Reidel, Dordrecht, pp. 133–145.
- Shmakin, A.B., 1998. The updated version of SPONSOR land surface scheme: PILPS-influenced improvements. *Glob. Planet. Change* 19 (1–4), 49–62.
- Slater, A.G., Schlosser, C.A., Desborough, C.E., Pitman, A.J., Henderson-Sellers, A., Robock, A., Vinnikov, K.Y., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P.M., de-Rosnay, P., Dickinson, R.E., Dai, Y.J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y.M., Habets, F., Kim, J., Koren, V., Kowalczyk, E.A., Nasonova, O.N., Noilhan, J., Schaake, S., Shmakin, A.B., Smirnova, T.G., Verseghy, D., Wetzel, P., Yue, X., Yang, Z.L., Zeng, Q., 2001. The representation of snow in land surface schemes: results from PILPS-2d. *J. Hydrometeorol.* 2, 7–25.
- Steele, M., Thomas, D., Rothrock, D., Martin, S., 1996. A simple model study of the Arctic Ocean freshwater balance, 1979–1985. *J. Geophys. Res.* 101c, 20833–20848.

- Storr, D., Golding, D.L., 1974. A preliminary water balance evaluation of an intensive snow survey in a mountainous watershed. *Advanced Concepts and Techniques in the Study of Snow and Ice Resources*. Nat. Acad. Sci., Washington, DC, pp. 294–303.
- Sud, Y.C., Mocko, D.M., 1999. New snow-physics to complement SSiB: Part I. Design and evaluation with ISLSCP initiative I datasets. *J. Meteorol. Soc. Jpn.* 77 (1B), 335–348.
- Sun, S., Xue, Y., 2001. Implementing a new snow scheme in Simplified Simple Biosphere Model (SSiB). *Adv. Atmos. Sci.* 18, 335–354.
- Takata, K., Kimoto, M., 2000. A numerical study on the impact of soil freezing on the continental-scale seasonal cycle. *J. Meteorol. Soc. Jpn.* 78 (3), 199–221.
- Tilley, J.S., Lynch, A.H., 1998. On the applicability of current land surface schemes for Arctic tundra: an intercomparison study. *J. Geophys. Res.* 103 (D22), 29051–29063.
- TVA, 1972. Heat and mass transfer between a water surface and the atmosphere. *Water Resources Report No. 0-6803 14*. Tennessee Valley Authority.
- van den Hurk, B.J.J.M., Viterbo, P., Beljaars, A.C.M., Betts, A.K., 2000. Offline validation of the ERA40 land surface scheme. *ECMWF Tech. Memo.* 295, 42 pp.
- Viterbo, P., Betts, A.K., 1999. Impact on ECMWF forecasts of changes to the albedo of the boreal forests in the presence of snow. *J. Geophys. Res. Atmos.* 104 (D22), 27803–27810.
- Viterbo, P., Beljaars, A., Mahfouf, J., Teixeira, J., 1999. The representation of soil moisture freezing and its impact on the stable boundary layer. *Q. J. R. Meteorol. Soc.* 125, 2401–2426.
- WCRP, 1996. Report of the fourth session of the WCRP ACSYS scientific steering group, Toronto, Canada, October 11–14, 1995, WCRP informal report N10.
- WCRP, 1999a. Report of the GEWEX/ACSYS workshop on cold regions hydrological modelling, Quebec City, Canada, 25–27 August 1998, WCRP Informal Report No. 13/1999.
- WCRP, 1999b. Report on the hydrology models intercomparison planning meeting, Koblenz, Germany, 27–29 March 1999, WCRP Informal Report No. 12/1999.
- Weatherly, J.W., Walsh, J.E., 1996. The effects of precipitation and river runoff in a coupled ice-ocean model of the Arctic. *Clim. Dyn.* 12 (11), 785–798.
- Weaver, A.J., Sarachik, E.S., Marotze, J., 1991. Fresh-water flux forcing of decadal and interdecadal oceanic variability. *Nature* 353 (6347), 836–838.
- Woo, M.K., 1982. Snow hydrology in the high Arctic, presented at . The Western Snow Conference, Reno, Nevada, April, 20–23.
- Woo, M.K., 1988. Wetland runoff regime in northern Canada. *Permafrost: Fifth International Conference 1, Proceedings, International Conference on Permafrost*, 644–649.
- Woo, M.K., Steer, P., 1983. Slope hydrology as influenced by thawing of the active layer, Resolute, N.W.T. *Can. J. Earth Sci.* 20, 978–986.
- Woo, M.K., Winter, T.C., 1993. The role of permafrost and seasonal frost in the hydrology of northern wetlands in North America. *J. Hydrol.* 141, 5–31.
- Wood, E.F., Lettenmaier, D.P., Liang, X., Lohmann, D., Boone, A., Change, S., Chen, F., Dai, Y., Dickinson, R., Duan, Q., Ek, M., Gusev, Y., Habets, F., Irannejad, P., Koster, R., Mitchell, K., Nasonova, O., Noilhan, J., Schaake, J., Schlosser, A., Shao, Y., Shmakin, A., Verseghy, D., Warrach, K., Wetzel, P., Xue, Y., Yang, Z., Zeng, Q., 1998. The project for intercomparison of land-surface parameterization schemes (PILPS) phase 2(c) Red–Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. *Glob. Planet. Change* 19, 115–135.